Development of Uncharged Galactosyltransferase Inhibitors Chemical Tools for Applications in Cells

Jiang, Jingqian

Awarding institution:
King's College London

Development of Uncharged Galactosyltransferase Inhibitors: Chemical Tools for Applications in Cells

By

Jingqian Jiang

Supervisor: Dr Gerd K. Wagner

A thesis submitted to the Faculty of Natural Mathematics and Science

at King's College London

December 2016

Abstract

β-1,4-Galactosyltransferases (β-1,4-GalTs) catalyse the transfer of D-galactose from a uridine diphosphate galactose (UDP-Gal) donor to an *N*-acetylglucosamine (*N*-GlcNAc) or glucose (Glc) acceptor, forming a β-1,4-glycosidic linkage. β-1,4-GalTs are required for the formation of important glycan epitopes, such as terminal tetrasaccharide Sialyl Lewis X (sLe^x), which is present in P-selectin glycoprotein ligand 1 (PSGL-1) and other cell adhesion molecules. Therefore, small molecular β-1,4-GalT inhibitors are of great interest as chemical tool compounds to study sLe^x- and PSGL-1-dependent processes. A UDP-Gal derivative, 5-(5-formylthien-2-yl) UDP-galactose (5-FT UDP-Gal), has previously been described as a potent, broad-spectrum GalT inhibitor; however, the application of 5-FT UDP-Gal in cell assays is compromised by its limited stability and membrane permeability, due to the presence of pyrophosphate and sugar moieties. Therefore, the main aim of this thesis was to develop uncharged β-1,4-GalT inhibitors based on 5-FT UDP-Gal, but with more suitable properties for cellular applications.

Several approaches were explored to achieve this goal. In chapter 2, attempts to apply the pro-drug concept using phosphate esters of 5-FT UDP-Gal are described. A series of 5-substituted nucleoside derivatives derived from 5-FT UDP-Gal was also prepared. The inhibitory activities of these derivatives against β-1,4-GalT were assessed in biochemical assays. Direct comparison with the corresponding complete UDP-sugar derivatives allowed the identification of structural factors that determine activity. The effects of the most active nucleoside derivative and its ester prodrug were also investigated in a PSGL-1 expression assay.

Attempts to overcome the relative loss of activity from the absence of pyrophosphate and sugar moieties in nucleoside inhibitors using dynamic combinatorial chemistry are described in Chapter 3. A hydrazone dynamic combinatorial library (DCL) was generated from the most potent nucleoside fragment and a series of hydrazides to identify mimics of pyrophosphate and sugar moieties to develop potent inhibitors. A suitable hydrazide was identified from the library and the corresponding nucleoside derivatives were generated and evaluated in the biochemical assay as well as the PSGL-1 expression assay.

A known, substrate-based β-1,4-GalT inhibitor was prepared as a positive control in the DCL experiments. However, this *N*-GlcNAc derivative unexpectedly behaved as an acceptor substrate rather than an inhibitor in our phosphatase-coupled assay. These unexpected finding, including attempts to rationalise the discrepancy between these results and reported in previous literature describing this compound as a β-1,4-GalT inhibitor, are described in Chapter 4.

Table of Contents

List of Figures

Figure 90 Effect of the β-1,4-GalTs and phosphatase on the production of glycosylated product.

Figure 91 Proposed explanation for the differential results for substrate and inhibitor activity of compound 43. A: GTs acceptors; D: GTs donors; E: GTs; A^*: glycosylated acceptors; D^*: nucleoside diphosphate (NDP).

List of Schemes

List of Tables

Table 18 Gradient B utilised for UDP-Gal and compound 43

Acknowledgements

Firstly, I would like to deeply thank my first supervisor, Dr Gerd K. Wagner for the guidance and support in my academic work. I also appreciate his help and feedback during the preparation of my thesis. I wish to thank my second supervisor, Prof. Clive Page for his great support and valuable suggestions for my PhD project in the last four years.

My thanks go to King's College London for the studentship, which enabled me to carry out my PhD study.

I would also like to thank the EPSRC National Mass Spectrometry Centre (Swansea), for the recording of mass spectra.

Special thanks go to Dr. Varsha Kanabar-Raivadera and Dr. Simon Pitchford in the Institute of Pharmaceutical Science at King's, who evaluated the inhibitory activity of my compounds in cell assays. Special thanks go also to Dr Lauren Tedaldi for her guidance during the early stages of my project. She was always patient when training me in the biochemical assays. Thanks go also to Dr Niina Goos for providing me with LgtC enzyme. I want to thank Melissa Edmondson for training me in the use of the HPLC for my work. I would like to thank Dr Sarah Barry for her many useful suggestions during the research group meetings.

I also would like to sincerely thank all past and present members of the Wagner group: Sarah, Ferdinand, Alice, Yong, Camille for their help and guidance during my hours in the lab. I also wish to thank my colleagues Yongmin, Chen, Ivy, Louise, Cathy, who made me feel happy to work in the lab in the last several years.

Finally, I would like to thank my parents for supporting me to study abroad and encourage me to overcome the challenges encountered in work and life.

Abbreviations

Abbreviation	Meaning
^{o}C	Degree Celsius
µL	Microlitres
µM	Micromoles
5–FT UDP-Gal	5-formylthiophene-UDP-Gal
5–FT uridine	5-formylthiophene uridine
Arg	Arginine
br	Broad singlet
BSA	Bovine serum albumin
CAZy	Carbohydrate-active Enzymes
CE	Capillary Electrophoresis
CEL	Chicken egg-white lysozyme
CIP	Calf intenstinal alkanine phosphatase
CH_2Cl_2	Dichloromethane
$CHCl_3$	Chloroform
CMP-Sialic acid	Cytidine monophosphate Sialic acid
C.*jejuni*	*Campylobacter jejuni*
COSY	Correlation spectroscopy
Cys	Cysteine
DCC	Dynamic combinatorial chemistry
DCL	Dynamic combinatorial library
d	Doublet
dd	Doublet of doublet
DME	Dimethoxyethane
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
E	Enzyme
E.C.	Enzyme Commision
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid

ELISA	Enzyme-linked Immunosorbant assay
Equiv.	Equivalent
ESI	Electrospray ionization
F	Fluorophore
FACE	Fluorosore assisted Capillary Electrophoresis
FRET	Fluorescence Resonance Energy Transfer
Gal	Galactose
Gal-β1,4-GlcNAc	N-acetylactamine
GalTs	Galactosyltransferases
GDP	Guanosine diphosphate
GDP-Man	Guanosine diphosphate mannose
GDP-Fuc	Guanosine diphosphate Fucose
Glc	Glucose
GlcTs	Glucosyltransferases
GlcNAc	Glutamic acid
gly	Glycine
glu	Glucotamine acid
GTs	Glycosyltransferases
h	Hour
HCl	Hydrochloride acid
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HPLC	High Pressure Liquid Chromatography
HTS	High-through Screening
IC_{50}	Half maximal inhibitory concentration
IEC	Ion-exchange chromatography
IPA	Isopropanol
IPTG	Isopropalthio-β-galactoside
Ki	Inhibitory constant of competitive inhibitor
Km	Michaelis-menten constant
LC	Liquid chromatography
leu	Leucine
LG	Leaving group

LgtC	α-1,4-galactosyltransferases from *Neisseria meningitidis*
LOS	lipooligosaccharides
LPS	Lipopolysaccharides
LS	Lactose synthase
Lys	Lysine
m	Multiplet
M	Molar
Me	Methyl
MeCN	Acetonitrile
MeOH	Methanol
MHz	Mega Hertz
MOPS	3-(N-morpholino) propansulfonic acid
mg	Miligram
min	Minute
mL	Mililitre
mol	Moles
MS	Mass spectrometry
mU	Miliunits
MurG	GlcNAc transferase
MW	Molecule weight
NADH	Nicotinamide adenine dinuclotide (reduced form)
NaOH	Sodium hydroxide
NB	Nucleobase
NDP	Nucleoside diphosphate
nm	Nanometre
nM	Nanomolar
NMP	Nucleoside monophosphate
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
OGT	O-GlcNAc transferases
Ph	Phenyl
pNPβGal	P-nitrophenyl-β-galactoside

ppGalNAcTs	N-acetyl-α-galactosaminyltransferases
ppm	Parts per million
R^2	Regression
RET	Resonance Energy Transfer
Rf	Retention factor
RP	Reverse phase
RT	Retention time
s	Second
sLe^x	Sialyl Lewis X
S_N2	Biomolecular nucleophilic substitution
S_Ni	Internal nucleophilic substitution
STD	Standard deviation
TEAB	Triethylammonium bicarbonate
Tf	Triflate
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TPPTs	Triphenylphosphine trisulfonate
Trp	Tryptophan
U	Units
UDP	Uridine diphosphate
UDP-2F-Gal	Uridine diphosphate-2-fluoro-galactose
UDP-Gal	Uridine diphosphate galactose
UDP-GalNAc	Uridine diphosphate N-acetylgalactose
UDP-Glc	Uridine diphosphate glucose
UDP-GlcA	Uridine diphosphate glucoronic acid
UDP-GlcNAc	Uridine diphosphate N-acetylglucosamine
UDP-sugar	Uridine diphosphate sugar
UDP-Xyl	Uridine diphosphate xylose
UMP	Uridine monophosphate
UV	Ultra Violet

Parts of this thesis have already been published in scientific journals:

Uncharged nucleoside inhibitors of β-1,4-galactosyltransferase with activity in cells. J. Jiang, V. Kanabar, B. Padilla, F. Man, S. C. Pitchford, C. P. Page, G. K. Wagner. *Chem. Commun.* **2016**, *52*, 3955-3958.

Based-modified UDP-sugars reduce cell surface levels of P-selectin glycoprotein 1 (PSGL-1) on IL-1β-stimulated human monocytes. V. Kanabar, L. Tedaldi, J. Jiang, X. Nie, I. Panina, K. Descroix, F. Man, S. C. Pitchford, C. P. Page, G. K. Wagner. *Glycobiology* **2016**, *26*, 1059-1071.

A practical glycosyltransferase assay for the identification of new inhibitor chemotypes. L. Tedaldi, A. Evitt, N. Goos, J. Jiang, G. K. Wagner. *Med. Chem. Commun.* **2014**, *5*, 1193-1210

CHAPTER 1

General introduction and objectives

1.1. Glycosyltransferases

Glycosyltransferases (GTs) are a large family of enzymes which are involved in the biosynthesis of glycoconjugates and glycans.[2] GTs catalyse the transfer of a sugar moiety from a glycosyl donor to acceptor molecules forming glycosidic bonds. Typical acceptors may contain saccharides, lipids, proteins and glycoconjugates.[1]

1.1.1. Mechanism of GTs

GTs elongate the glycan chains by transferring sugar moieties to acceptors, forming the glycosidic linkage. In principle, two stereochemical results are possible for sugar transferred reactions. The configuration of anomeric center in the product is either retained or inverted with respect to the sugar nucleotide donor.[3] (**Figure 1**).

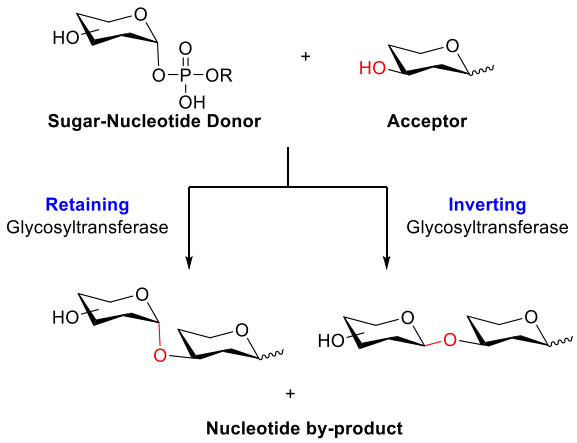


Figure 1 Retaining and inverting glycosylated reactions catalysed by GTs

The mechanism for inverting GTs is well established.[1] The inverting glycosyltransferases utilise a S_{N-2} like direct-displacement reaction and lead to an inverted anomeric configuration. An amino side chain at the active site of enzyme serves as a base catalyst to de-protonate the hydroxyl group of the acceptor, facilitating its nucleophile attack towards the donor. The reaction proceeds *via* the formation of an oxocarbenium ion transition-state, which was followed by the departure of the nucleotide leaving group. (**Figure 2**).

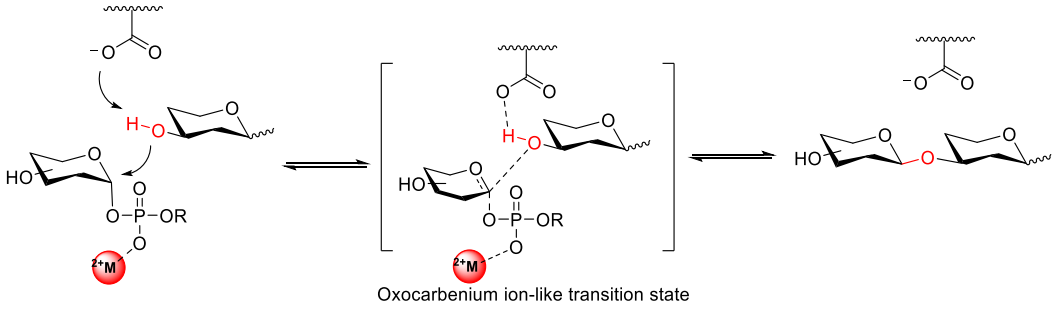


Figure 2 Mechanism of inverting enzyme catalysed transfer reaction (modified from reference[2])

Compared to the inverting GTs, the mechanism of retaining enzyme-catalysed reaction is still unclear. It is proposed that the retaining transfer processes a double displacement involving a covalently bound glycosyl-enzyme intermediate species (**Figure 3**). The double displacement mechanism was supported by the identification of intermediates with the glycosyl donor covalently linked to the enzyme.[4, 5] A divalent cation, like Mn^{2+}, or a suitably positioned positively charged side chain or a helix dipole acts as a Lewis acid for the donor phosphate. The leaving nucleotide group acts as a base catalyst to activate the hydroxyl group of the acceptor for nucleophilic attack.

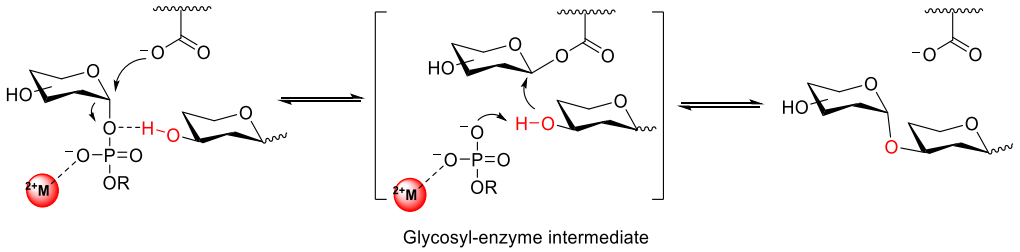


Figure 3 Proposed mechanism of retaining enzyme catalysed transfer reaction. (modified from reference[2])

A S_Ni-like mechanism was also proposed for the retaining GTs reaction, in which the acceptor is thought to attack the donor on the same face as the nucleotide leaving group, leading to retain the anomeric configuration of the glycosidic linkage.[5] (**Figure 4**)

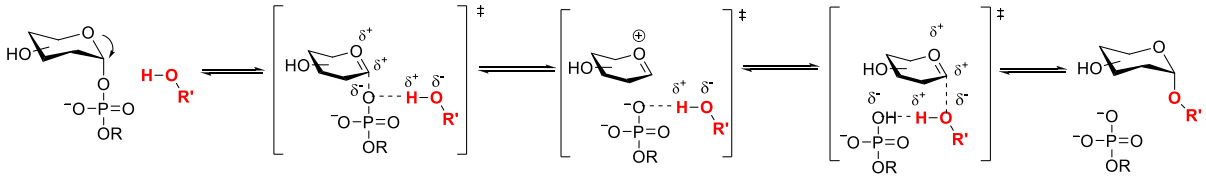


Figure 4 S_Ni-like mechanism of retaining enzyme catalysed reaction. (modified from reference[6])

1.1.2. Classification of GTs

Traditionally, GTs are classified according to their amino acid sequence similarities, the stereochemical outcomes of the reactions and their three-dimensional structures.[1] Until now, 94 GT families are classified in the Carbohydrate Active Enzymes Database (CAZy) based on the amino acid sequence similarities.[7] Structurally, a majority of the GTs can be classified into two classes, GT-A and GT-B folds, based on their three-dimension crystal structures.[1] GT-A fold is made of a single Rossman fold consisting of an open twisted β-sheets surrounded by α-helices. (**Figure 5a**) The GT-A fold was first identified when the crystal

structure of SpsA from *B. subtilis*, a member of the Leloir pathway enzyme, was described.[8] Most GT-A folded GTs have Asp-x-Asp (DXD) motif in the centre of active sites that coordinate with pyrophosphate moiety of nucleotide donor via metal cations, like Mn^{2+} and Mg^{2+}. The acceptor binding sites are generally made of two flexible loops that undergo conformational changes following binding of the donor.[9]

GT-B fold enzyme contains a two-domain structure with a Rossmann-like fold in either domain. (**Figure 5b**) The predicted active sites of GT-B folded enzymes are located between the two Rossmann folds. Compared with GT-A fold enzymes, the joined helices of the subdomain of Rossman folds create binding sites for acceptors. The GT-B folds GTs include most prokaryotic enzymes that produce secondary metabolites and important bacterial cell wall precursors.[10]

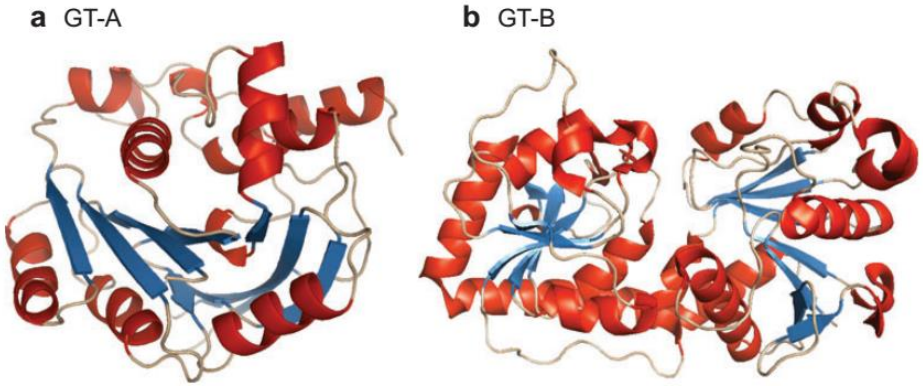


Figure 5 (a) Three-dimensional crystal structure of the enzyme SpsA from *Bacillus subtilus*, illustrating GT-A folds. (b) Bacteriophage T4 β-glucosyltransferase, pdb 1jg7 illustrating GT-B fold. (β-sheets are in blue, α-helices in red, strands in grey) (Figures taken from reference[1])

A GT-C fold enzyme family has been proposed and it exhibits many similarities with GT-A fold but utilizes lipid-linked donors. Only two crystal structures have been reported.[11, 12]

1.1.3. Donor and acceptor specificity of GTs

Many GTs are Leloir enzymes that utilise sugar nucleotides as their glycosyl donors. [13] These compounds are made of a nucleoside and a sugar moiety linked *via* a pyrophosphate bond. The highly energetic bond between pyrophosphate moiety and the anomeric centre of sugar ring activate the sugar for the transfer reaction. Until now, there are only nine sugar nucleotides reported as natural donors, including UDP-Glc, UDP-Gal, UDP-GlcNAc, UDP-Xyl, UDP-GlcA, GDP-Man, GDP-Fuc and CMP-Sialic acid. The amount of sugar nucleotide donors is more than in non-mammalian organisms, which offer an opportunity to target GTs to treat diseases, like bacterial infection, with little effects on human.

GTs exhibit high specificity towards both donor and acceptor substrates. However, some GTs also can tolerate different donor and acceptor substrates.[14-16] They have been utilised as tools for the enzymatic production of synthetically challenging oligosaccharides.[17] GTs also can exhibit donor flexibility as some enzymes accept various non-natural sugar nucleotide donors.[18, 19]

1.1.4. GTs in biological systems

Products of GTs catalysed reactions, glycans and glycoconjugates, mediate many fundamental biological processes. They can serve as energy sources[20], locate on the cell wall where they act as receptors in recognition processes, like cell signalling and cellular adhesion[21], as well as provide receptors for bacteria, hormones and virus[22]. GTs are important for the biosynthesis of *N*-/*O*-linked glycoproteins and glycolipids in both prokaryotic and eukaryotic organisms.[23, 24] Eukaryotic GTs are primarily located in the endoplasmic reticulum (ER) and Golgi apparatus and products, glycolipids as well as glycoproteins, are

mainly on cell surfaces.[25] These type II transmembrane proteins can serve as receptors for cell adhesion. In prokaryotic cells, GTs catalysed reactions occurs in the cytoplasm and periplasm.[26] Glycans in prokaryotic cells are essential as some cellular roles. Lipopolysaccharides (LPS) can act as attachment receptor for cellular interactions and are involved in processes of modulation of immune system, as well as acting as virulence factor for the infection of the host.[27] Peptidoglycans, which is formed of linear glycan strands cross-linked by short peptide, are determinant of the cell wall and essential for cell shape and protects the bacteria from turgidity.[28]

The saccharides transferred reactions are involved in diseases. Changes in glycosylation will cause under- or over expression of glycans. For example, cancer cells usually display glycans at different levels or with fundamentally different structures from the normal cells. The increasing activity of *N*-acetylglucosaminyltransferase V (GnT-V, the enzyme that leads to β1,6GlcNAc branching) leads to an increase in the size and branching of *N*-linked glycans. These extensively β1,6GlcNAc branched glycans creat additional sites for terminal sialic acid residues on the cell surface, leading to an increasing activity of sialytransferases in cancerous cells.[29] Increased expression of Lewis antigen is a common feature of many epithelian cancers.[30] Fucosyltransferase FucT I-VIII, which transfer fucose in α-1,2,3,4, or 6-linkage to galactose or GlcNAc residues, are involved in the formation of sialyl Lewis A. The activities of these enzymes were observed upregulation in colorectal cancers. The increasing activity of FucT I-VIII leads to the over formation of terminal glycans, sialyl Lewis X and Lewis Y.[31]

1.2. β-1,4-galactosyltransferases

1.2.1. Mechanism of β-1,4-GalTs

Amongst the most extensively studied and best characterized GTs are β-1,4-galactosyltransferase (β-1,4-GalTs).[32] β-1,4-GalTs are involved in transferring galactose from uridine diphosphate galactose (UDP-Gal) to the 4-OH of terminal *N*-acetylglucosamine (GlcNAc), Glucose (Glc) or Xylose (Xyl) residues, forming β-1,4 glycosidic linkage (**Figure 6**). When as a single protein, β-1,4-GalTs catalyse the transfer of galactose from UDP-Gal to *N*-acetylglucosamine (GlcNAc) based acceptors to form *N*-Acetyllactosamine. Nevertheless, when it interacts with the calcium binding protein α-lactalbumin (LA), β-1,4-GalTs form a lactose synthease complex and the binding of LA changes the acceptor specificity of β-1,4-GalTs. This complex catalyses the transfer of galactose to glucose to form lactose.[33]

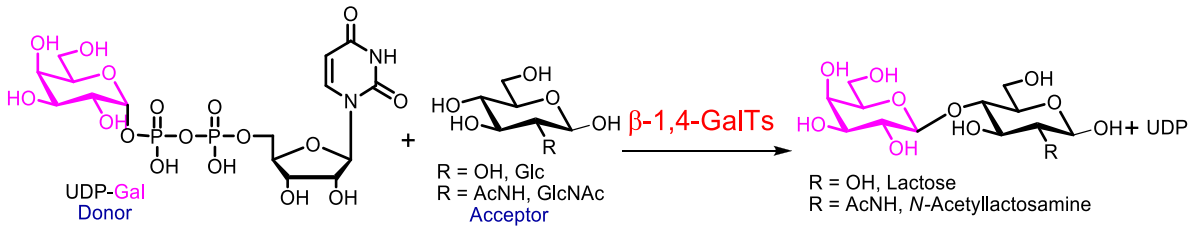


Figure 6 Generalized β-1,4-galactosyltrasnferase reactions

The resolution of three-dimensional crystal structures of β-1,4-GalTs has facilitated investigations into the mechanism of the transfer reaction catalysed by β-1,4-GalTs.[34, 35] The catalytic pocket is located between the two domains, which contains two flexible loops. The

two flexible loops, a long one (residues 345-365) and a short one (residues 313-316), play crucial roles in catalytic cycle. (**Figure 7**)

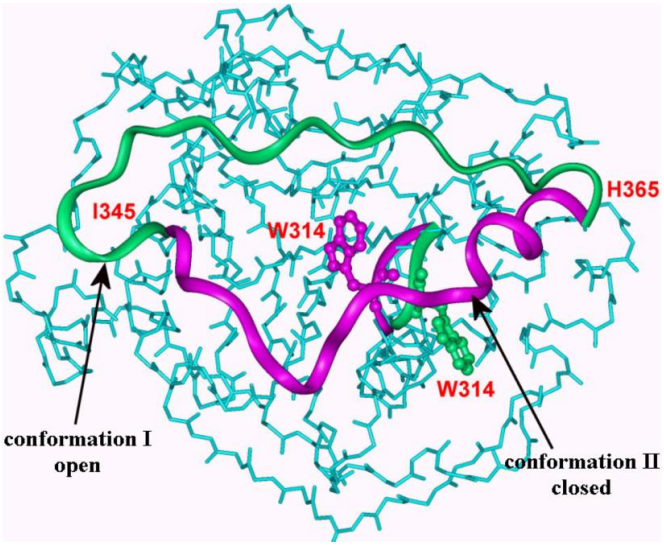


Figure 7 The long flexible loop, residue IIe345-His365, and the short flexible loop, residue Gly313-Gly316, in the open (green) and closed (purple) conformations. Residue Trp314 is facing away from the open conformation and moves to coordinate the metal ion, UDP-Gal and acceptor in the closed conformation. (Figure taken from reference[36])

In the substrate-free state, the pocket is in an open conformation and loops are both oriented such that important residues participating in catalysis are distant from the pocket. In the catalytic cycle, the metal ion Mn^{2+} binds with the enzyme, which is followed by the binding of donor UDP-Gal. Upon these binding, the flexible loops in catalytic domain undergo a marked conformational change, from an open conformation to a closed conformation, creating the acceptor binding site. Tryptophan 314 on the short loop plays a crucial role in the catalytic cycle, interacting with both UDP-Gal and acceptor during the closed conformation.[37] After completion of the transfer of Gal to acceptor, the glycosylated product is ejected. Then the enzyme reverts to the original conformation, releases the left nucleotides and metal ion and repeats the catalytic cycle. (**Figure 8**)

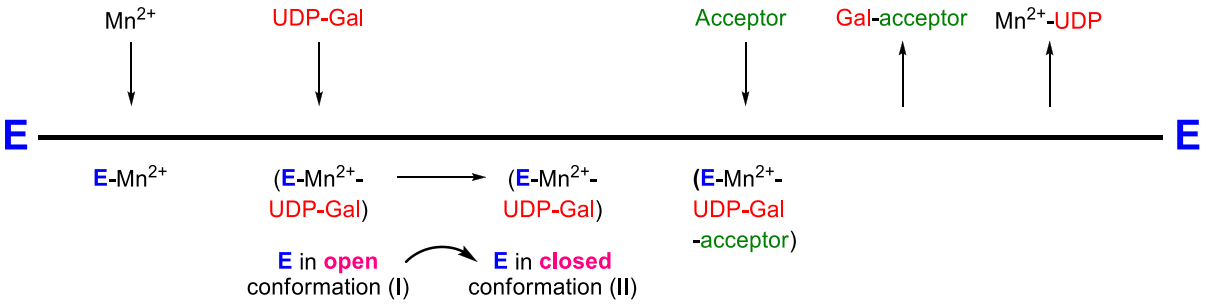


Figure 8 Mechanism of β-1,4-GalTs catalysed transfer reaction. (Figure taken from reference[34])

1.2.2. β-1,4-GalTs in biological systems

The first galactosyltransferase to be cloned was the β-1,4-GalT-I and it was the most thoroughly characterized.[38] β4GalT1 is a trans-Golgi resident, type II membrane-bound glycoprotein, which consists of catalytic domain, stem region, a single membrane-spanning region as well as a short cytoplasmic N-terminus. (**Figure 9**) To date, seven β-1,4 GalTs have been described. These proteins share an extensive homology, which can be used for clone the gene to encode six-members of the family. According to their percentage sequence identity at amino acid level relative to β4-GalT-I, which was the first identified β-1,4 GalT, they are named from β4GalT-II to β4GalT-VII.[35] β–1,4-GalT-I to -VI have several conserved regions.[39, 40] These conserved regions may provide the binding site of sugar donor or acceptor, which means they may share similar enzymatic activity. Seven GalTs can use UDP-Gal as sugar donor and Glc, GlcNAc or Xyl as acceptor. β4GalT-I, II, III, IV, V, VI have the similar characteristic of enzyme kinetics.

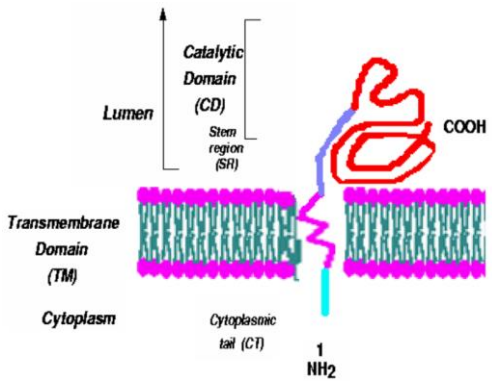


Figure 9 Schematic diagram of protein domains of β4Gal-T1. (Figure taken from reference[35])

It is known that β4GalTs are involved in the biosynthesis of different glycoconjugates and glycans. For example, like β4GalT-I, β4GalT-II, III, IV, V also have lactose synthase activity, transferring galactose from UDP-Gal to Glucose, forming lactose.[40] β4Gal-I can participate the biosynthesis of sialyl Lewix (sLe^x), a crucial carbohydrate determinant on PSGL-1, and blood group antigen.[41] β4Gal-I also can elongate *O*-fucose glycans found on several epidermal growth factor (EGF)-like protein domain.[42] β4GalT IV can prefer GlcNAc 6-*O*-sulfate (6SGN) as acceptor and involved in the biosynthesis of glycoproteins carrying a 6-*O*-sulfated *N*-acetyllactosamine moiety.[43] β4GalT-VII is involved in the biosynthesis of proteoglycans by catalysing transfer of Gal to glycosaminoglycan core.[44]

β-1,4-GalTs are expressed in all mammalian tissues and involved in many fundamental biological processes. For example, the glycosylated products of β-1,4-GalTs act as receptors on the eukaryotic cell surface and are involved in cellular interactions by binding to particular extracellular glycoside substrates. Thus, β-1,4-GalTs are involved in the neurite extension, sperm-egg interaction, cell adhesion and migration.[45-47] β-1,4-GalTs are also expressed in bacteria and involved in the relating biological processes. For example, lipopolysaccharide (LPS) structures contain *O*-polysaccharide (*O*-antigenic) side chains linked to a core

polysaccharide and lipid A.[48] LPS are located in the outer membrane of bacterial cell surface and are crucial for the interaction between bacteria and their environment.[48] *Shigella* is a Gram-negative human pathogen which causes diarrhoea in humans and β-1,4-GalTs are involved in the biosynthesis of the repeating oligosaccharide units of LPSs of *Shigella*.[48] Deletion of β-1,4-GalTs producing in mice leads to decreasing activity of β-1,4-GalTs in most tissues, decreasing of fertility and differentiation of epithelial cells retardation.[49] β-1,4-GalTs are involved in embryo implantation and the growth and adhesion of cancer cells.[50] [51] Moreover, increase and decrease of β−1,4-GalTs activity are associated with disease states. The level of cell surface β-1,4-GalTs is increased in a variety of metastatic murine and human cell lines. Decreasing surface β-1,4-GalT expression on highly metastatic cells to the level characteristic of low metastatic cells could reduce their invasive behaviour in vitro and metastatic activity in vivo.[52]

1.2. P-selectin glycoprotein ligand-1 (PSGL-1)

Leukocyte recruitment is an important immunological process which recruits free leukocytes from intravascular compartment to transfer into sites of inflammation. It predominantly occurs in inflamed microvessels and in high endothelial venules of secondary lymphoid organs.[41] It proceeds along capture of free flowing leukocytes and rolling of leukocytes steps. The rolling step, which is considered as an important step for the successful recruitment of leukocytes into tissues, is mediated by a family of glycoproteins, selectins. In the process, these selectins recognize and bind with the crucial carbohydrate determinants with selectin ligands (**Figure 10**). Among these ligands, P-selectin glycoprotein ligand-1 (PSGL-1) is the

dominant P-selectin ligand during inflammation. In in-vivo experiments, PSGL-1 deficient mice were used for intravital microscopy experiments in postcapillary venules of the cremaster muscle. It was shown that leucocyte rolling was significantly decreased 30 min after trauma.[53] Likewise, PSGL-1 also can mediate rolling of leukocytes on E-selectin under flow as well as act as the ligand towards L-selectin. This was supported by the experiment that in PSGL-1 deficient mice, L-seletin dependent leukocytes rolling was completely absent in acute inflammation and the fluid-phase E-selectin bound to nearly 70% fewer sites than on the normal mice with PSGL-1.[54]

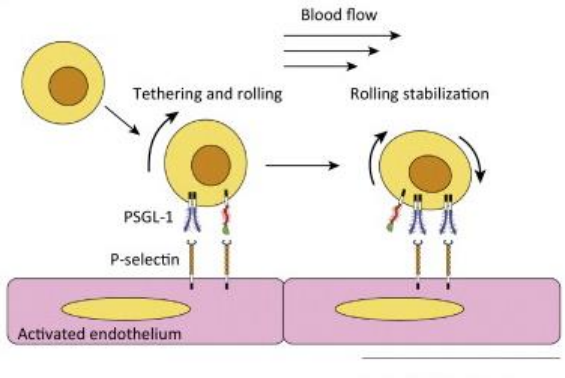


Figure 10 P-selectin mediated cell interactions with the vascular endothelium during inflammation. (Figure taken from reference[55])

PSGL-1 belongs to a growing number of glycoproteins and its function requires proper post-translational glycosylation, which is catalysed by relating glycosyltransferases. The synthesis of core decorated *O*-glycans carrying sialyl Lewix epitope is illustrated in **Figure 11**. Several glycosyltransferases have been identified to participate in the biosynthesis pathway, such as core 2 β-1,6-GlcNAc transferase, β-1,4-GalT-I and -IV, Fucosyltransferase-VII and -IV and α–2,3-sialyltransferase-IV, etc. These glycosyltransferases are all type-II transmembrane proteins situated in the Golgi apparatus. They transfer monosaccharides from

activated sugar donors such as UDP-GlcNAc, UDP-Gal, UDP-fucose and CMP-sialic acid to their respective glycoconjugate acceptor.

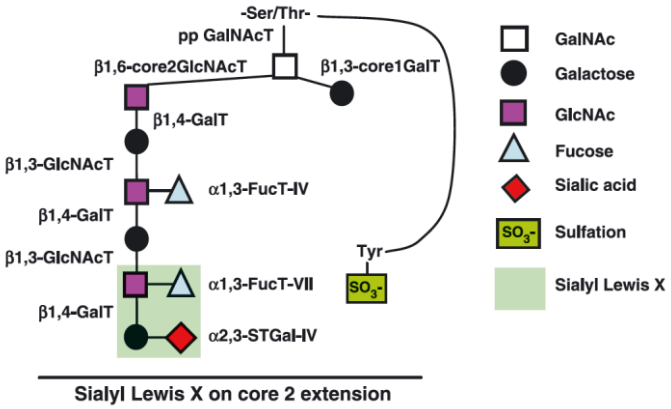


Figure 11 Biosynthesis pathway for the synthesis of core decorated *O*-glycans carrying the sLe^x determinant. (Figure taken from reference[41])

It is known that β-1, 4-GalT-I and β-1, 4-Gal-IV are involved in the synthesis of functional selectin ligand PSGL-1, catalysing the addition of UDP-Gal to GlcNAc in the process of forming sLe^x. Therefore β-1, 4-GalTs play key roles in the processes of inflammation mediated by PSGL-1. A β-1,4-GalT-deficient mice model was used to investigate the contribution of the enzyme for selectin ligand function.[56] The result showed that more than 80% of the core 2 *O*-glycans on the leukocyte membrane glycoproteins of these mice lacked β-1,4-linkage galactose residues. The soluble P-selectin binding to neutrophils and monocytes was strongly reduced. The biosynthesis of selectin ligand was significantly impaired. As for the aspects of acute and chronic inflammatory responses, zymosan-induced inflammation is a model in which acute inflammation is mainly mediated by E-and P-selectins. Zymosan-induced neutrophil recruitment was fully reduced and ear swelling was significantly suppressed in β-1,4-GalT-I deficient mice.[56] Contact hypersensitivity (CHS) and delayed-type hypersensitivity (DTH) responses are used as a model of chronic inflammatory

diseases. The experiment result showed that the CHS and DTH responses were partially suppressed in enzyme deficient mice. Due to the amount of P-selectin ligands deficiency in mice lacking β-1,4-GalT, neutrophil trafficking to sites of inflammation was impaired. [56]

1.3. Inhibitors of β-1,4-galactosyltransferases

β-1,4-GalTs are involved in the biosynthesis of complex glycans and their glycoconjugates that participate in many fundamental biological processes, as illustrated in the previous chapter. Since many of these biological products are essential for living organisms, inhibitors of β-1,4-GalTs are of great interest as chemical tools, not only for mechanistic and structural studies with recombinant enzymes, but also to investigate the role of β-1,4-GalTs in glycan biosynthesis. For example, a UDP-Gal derivative with a substituent in 6-OH of sugar ring was developed on the basis of 3D docking simulation to investigate the functional role of Trp310 residue of the small loop in the enzyme active site.[57] Therefore, the development of enzyme inhibitors attracts considerable interest and are considered to become potential lead compounds for drug discovery.

In the transfer reaction catalysed by β−1,4-GalTs, the reaction is proposed via a transition-state illustrated in **Figure 12**. The pyrophosphate moiety in UDP-Gal is suggested to form a six-member ring in its complex with an essential di-valent manganese ion. The side chain in enzyme active site serves as base to deprotonate the hydroxyl group of acceptor sugar, facilitating its attack ability towards the sugar moiety of donor.

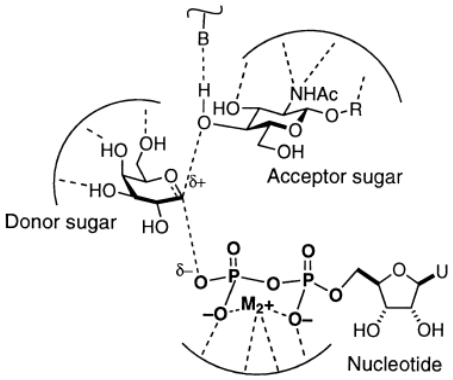


Figure 12 A possible transition-state structure of the β-1,4-galactosyl- transferase reaction. (Figure taken from reference[58])

Different strategies have been used to design potent β-1,4-GalTs inhibitors. To date, different types of β-1,4-GalTs inhibitors have been developed based on this proposed transition-state, including donor analogues, acceptor analogues and transition state analogues.

1.3.1. Donor analogues

The natural donor of β-1,4-GalTs, UDP-Gal, is composed of three moieties, galactosyl, pyrophosphate and uridine (**Figure 13**). Different approaches based on modifying sugar and phosphate moieties respectively have been used to design inhibitors of β-1,4-GalTs.

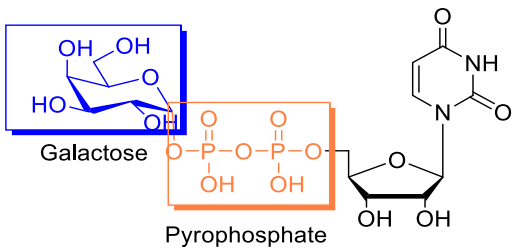


Figure 13 Design strategies for GalT inhibitors based on the UDP-Gal donor

1.3.1.1. Sugar analogues

Generally, the modifications of sugar analogues are made at the hydroxyl groups or the oxygen atom of the six-membered cyclic acetal. Hindsgaul and co-workers reported the chemical synthesis of uridine 5'-(5a-carba-α-D-galacopyranosyl) diphosphate, in which the ring oxygen of galatose residue was replaced by a methylene group.[59] (**Figure 14**) As the carbocyclic analogue of UDP-Gal, it acted as a competitive inhibitor towards β-1,4-GalTs with a K_i value of 58 μM, similar to the K_m value (25 μM) of UDP-Gal.

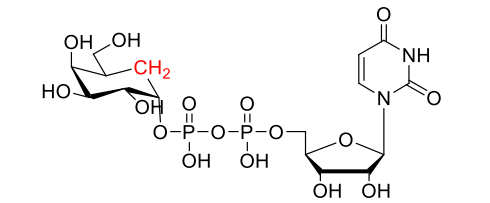


Figure 14 Carbocyclic analogue of UDP-Gal as the inhibitor against β-1,4-GalTs.

Wong and co-workers reported the preparation of uridine 5'-diphospho-(2-deoxy-2-fluoro) galactose (UDP-2F-Gal).[60] It was found to be a competitive inhibitor of β-1,4-GalTs with a K_i value of 149 μM. (**Figure 15**) Similarly, Hashimoto and co-workers also reported the sugar analogues, a series of mono-*O*-methyl analogues of UDP-Gal. Among these analogues, the 2-*O*-methyl analogues could inhibit β-1,4-GalTs. (**Figure 15**)

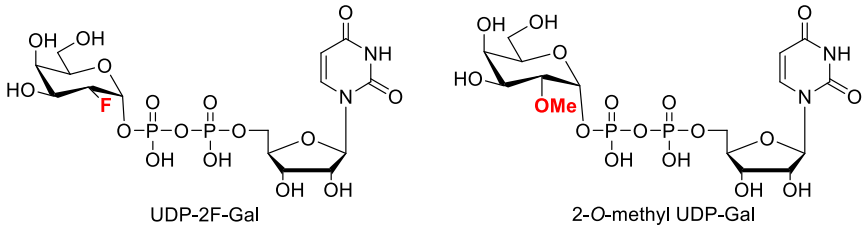


Figure 15 Sugar analogues with modifications on OH groups of galactose ring

Nishimura and co-workers reported the synthesis of uridine 5'-(6-amino-[2-(7-bromomethyl-2-naphthyl) methoxy]carbonyl)methoxy] acetyl-6-deoxy-α-D-galactopyranosyl) diphosphate according to 3D docking simulation.[61] (**Figure 16**) By investigating the interaction between compound and enzyme, it was suggested the Trp310 of β-1,4-GalTs can be selectively modified with the naphthylmethyl group of this compound. According to this results, a series of analogues were synthesized, among which compound containing naphthyl group was the strongest competitive inhibitor (K_i = 1.86 μM) against UDP-Gal (K_m = 4.91 μM).

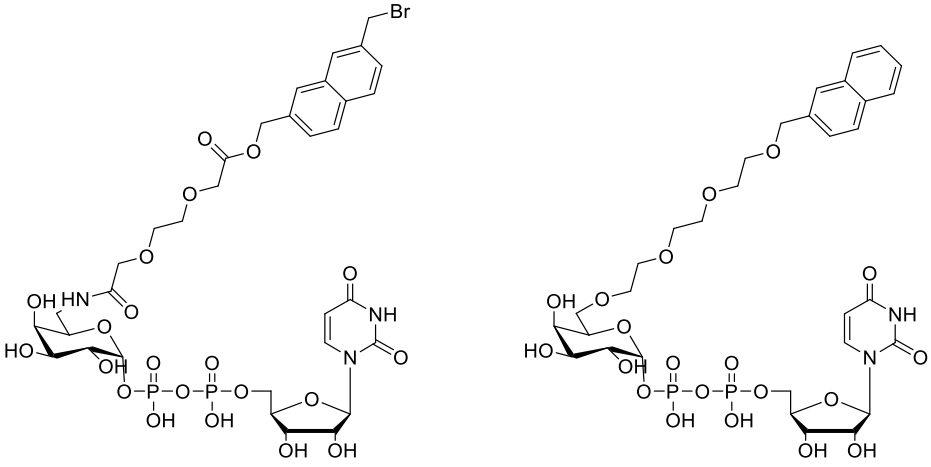


Figure 16 Inhibitors of β-1,4-GalTs with modification on 6'-OH *via* 3D docking simulation

1.3.1.2. Pyrophosphate analogues

The pyrophosphate group in UDP-Gal makes an important contribution to donor binding by coordinating to a di-valent manganese ion in the GalT active site. However, the pyrophosphate is quite unstable and due to its inherent negative charge, compounds containing pyrophosphate exhibit poor membrane permeability. Thus, pyrophosphate analogues were designed and synthesized by varying the diphosphate moiety.

Malonic, tartaric acid and monosaccharide linkages were utilized as pyrophosphate mimics to synthesis a series of compounds by Wong and his co-workers.[62] (**Figure 17**) It was hypothesised that the malonic and tartaric esters could form the six-member rings with Mn^{2+} which was similar as the Mn-pyrophosphate complex. The mono-sugar, like glucose and galactose, also could mimic this complex. The compound, in which glucose was used as pyrophosphate mimic, showed the inhibitory activity towards β-1,4-GalTs with a K_i of 119.6 μM.

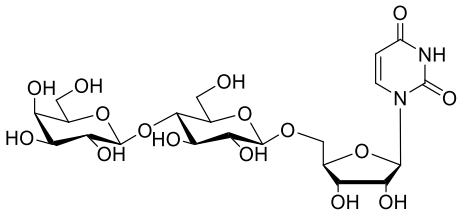


Figure 17 β-1,4-GalTs inhibitor containing monosaccharide as mimic of pyrophosphate

Matsuda and co-workers reported a series UDP-Gal analogues, in which pyrophosphate was replaced by alkene or alkane.[63] These compounds had the same number of atoms between the 5'-position of uridine and the anomeric position of galactose. With the introduction of an olefin and a hydroxyl group, a systematic structure-activity relationship could be conducted

in terms of the distance between the uridine and Gal and the relative orientation toward each other. (**Figure 18**). Yuasa and co-workers reported that the amino sugar moiety could serve as pyrophosphate due to chelation with manganese ion through the hinge motion.[64] Among the series of compounds, β-xylo-1-Uridine conjugate showed a moderate inhibition (45%). (**Figure 18**)

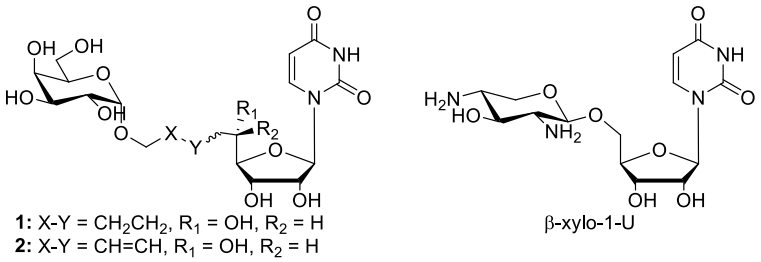


Figure 18 Acceptor mimics: alkane and alkene mimic pyrophosphate and β-xylo-1-U

1.3.2. Acceptor analogues

The modification of donor UDP-Gal has been used as a strategy to develop potential β-1,4-inhibitors. However, these inhibitors are likely to inhibit more than one GalT, since other GalTs also use UDP-Gal as donor substrate. This selectivity problem is common in the development of glycosyltransferase inhibitors. The β4GalT family consists of enzyme proteins that share sequence similarity, but they differ in tissue expression and possibly in acceptor substrate specificity. Therefore, consideration of the specificity of GalTs substrate, the acceptor, GlcNAc and glucose, analogues also can be developed as inhibitors of β-1, 4-GalT (**Figure 19**).

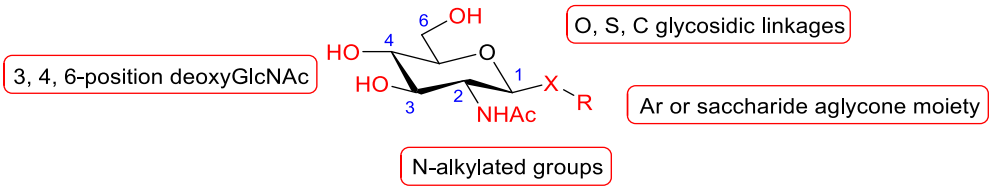


Figure 19 Design of GlcNAc derivatives as inhibitors towards β-1, 4-GalTs

A series analogues of GlcNAc was designed and synthesized by Brockhausen and co-workers.[65] Modifications were made at the 3-, 4- and 6-positions of the sugar ring of the acceptor, in the nature of the glycosidic linkage, in the aglycone moiety and in the acetamido group. Derivatives containing a 2-naphthyl aglycone were found to inhibit the transfer of Gal to substrate (**Figure 20**).

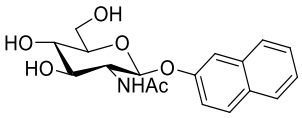


Figure 20 GlcNAcβ-(2-naphthyl), acceptor analogue with 2-naphthyl aglycone

A sugar analogue of GlcNAc, 2-naphthyl 2-butanamido-2-deoxy-1-β-D-glucopyranoside (**Figure 21**) was found to a be a specific inhibitor of β-1,4-GalTs in cell homogenates (human airway cells; H292, human colonic cancer cells, $CaCo^{-2}$ and mouse lymphocytic cells (MOPC)). Brown and co-workers reported an acceptor analogue, per-*O*-acetylated GlcNAc-β1,3-Gal-β-*O*-naphthalenemethanol, could inhibit sLe^x formation in U937 monocytic leukemia cells, suggesting its inhibitory activity in vivo.[66] (**Figure 21**)

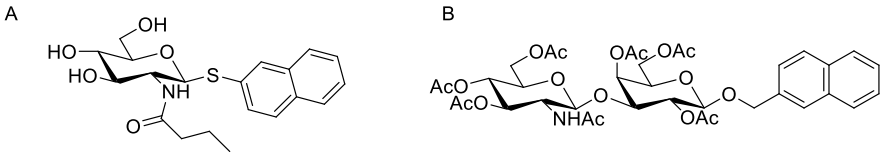


Figure 21 A: GlcNAc analogue 2-naphthyl 2-butanamido-2-deoxy-1-β-D-glucopyranoside; B: Acceptor analogue per-*O*-acetylated GlcNAcβ1-3Galβ-*O*-naphthalenemethanol.

1.3.3. Transition state analogues

Kajihara and co-workers reported the design of tricomponent bisubstrate analogues based on the model of S_N2-like transition-state.[67] (**Figure 22**) This compound showed remarkable inhibitory activity with a K_i of 1.35 μM towards β-1,4-GalTs against UDP-Gal with a K_i of 3.3 μM.

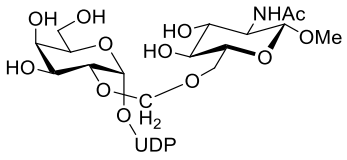


Figure 22 Tricomponent bisubstrate analogues towards β-1,4-GalTs

1.4. Strategies to improve cell membrane penetration of enzyme inhibitors

GTs play key roles in many biological processes and are potential drug targets. GT inhibitors are therefore sought after not only as chemical tools-, but also as potential lead compounds for drug discovery. Some GTs inhibitors are active in pharmacological assays. For example, a steroid-based sialyltransferase (STs) inhibitor AL10 is membrane-permeant and can effectively decrease total sialylation on the cell surface. Compound AL10 inhibits adhesion,

migration, actin polymerization and invasion of human lung cells which overexpress α-2,3-ST.[68] However, conventional donor and substrate analogues are frequently less suitable for drug development and, more generally, for cellular applications due to their limited membrane permeability and chemical stability. Therefore, strategies are required to improve membrane penetration of GT inhibitors for their application in cell assays.

Cell membrane surrounds the cytoplasm of the cell and in eukaryote cells, physically separating the intracellular organelles from the extracellular environment. Cell membrane is composed of a mix of proteins and phospholipids. Phospholipids form a lipid bilayer in which their electrically charged and hydrophilic head areas spontaneously arrange to face the aqueous cytosol and the extracellular fluid, while the hydrophobic tail area face towards each other, away from the cytosol. (**Figure 23**) Cell membranes are selectively permeable, allowing some substances to pass through while restricting the transport of others, which is essential for providing nutrients, eliminating waste as well as preventing unwanted molecules into cells. Due to the structure of the phospholipid bilayer, cell membranes are permeable through passive diffusion for uncharged and non-polar molecules, such as steroids, but less to for charged, polar molecules, such as sugar-nucleotides.

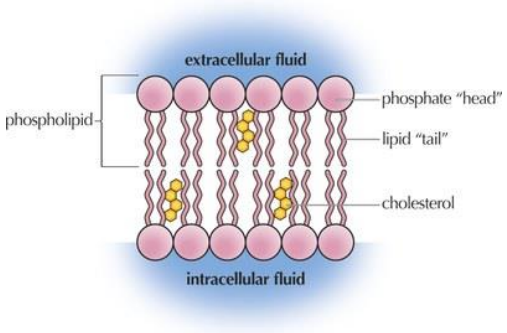


Figure 23 Structure of phospholipids bilayer. (Figure taken from reference[69])

Various strategies have been developed to improve the membrane permeability of polar compounds.[70] One of the most common approaches is to mask or modify the polar moiety itself by generating suitable pro-drugs. It has been demonstrated that protection of polar groups, like amides, phosphate and sulfones by lipophilic substituents leads to membrane-permeant compounds.[71-73] For example, modification of amide NH groups by N-methylation is assumed to contribute to cell membrane permeability. The cyclo-[Leu_4-Pro-Tyr] hexapeptide (**Figure 24**) displayed poor membrane permeability due to the polar amide groups. As expected, the permethylated cyclic peptide has much higher permeability than the non-methylated compound in RRCK cells.[74] Esterification is carried out to protect the polar hydroxyl groups and the protecting groups can removed by intercellular esterases after up-taken into cells. As the di-ester prodrug of nucleoside PSI-6130, compound RG7128 (**Figure 24**) was modified by esterification of hydroxyl groups which can be removed after up-taken into cells. Compared with the parent compound, it exhibited better membrane permeability and bioavailability.[75] The removal or replacement of polar functional groups can also improve the ability of compounds to cross lipid bilayers. For example, the benzamidine structure exhibits anti thrombotic activity. The amidine group in this structure influences the effective absorption due to its strong basicity. Thus, the compound PRO 3112, in which the amidine group is replaced by an isoquinoline group, exhibits reduced basicity, leading to increasing absorption.[76] (**Figure 24**)

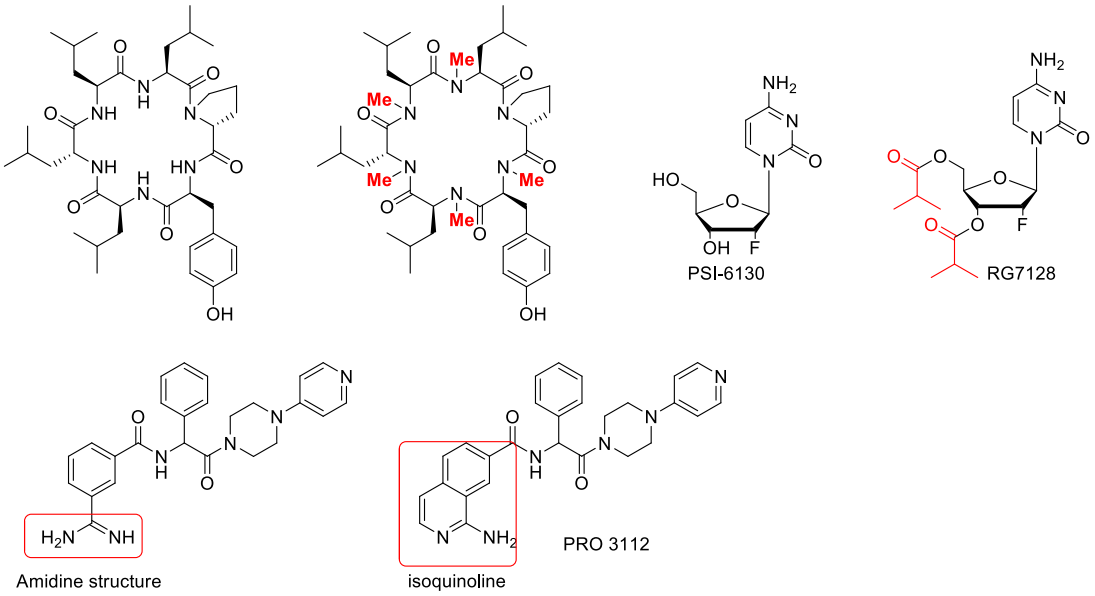


Figure 24 Cyclic hexapeptide and its permethylated analogue (left side); Nucleoside PSI-6130 as well as its ester prodrug. (right side); benzamidine structure and the analogue PRO-3112 containing isoquinoline structure. (bottom)

A number of studies have investigated how the formation of intermolecular hydrogen bond (IMHB) influences physicochemical properties and cell permeability. For example, in the study of the hexapeptide model system, the compound, cyclo-[D-Leu-D-Leu-Leu-D-Leu-Pro-Tyr], was found being the most permeable across a parallel artificial membrane (PAMPA), in which four hydrogen bonds were formed between amide NH and carbonyl O atom.[77] The influence of IMHB on compound properties was also demonstrated by stereoisomeric *Trypanosoma cruzi* growth inhibitors.[78] Computational chemistry and NMR spectroscopy results were used to identify the formation of the intramolecular NH-NR_3 bond in the compounds displaying higher cell permeability. (**Figure 25**)

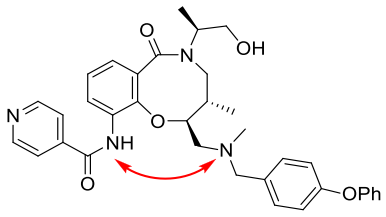


Figure 25 Structure of *Trypanosoma cruzi* growth inhibitors and the intramolecular NH-NR_3 bond in the compounds leads to higher cell permeability.

Macrocyclization has also been investigated as an alternative approach for improving the membrane penetration of drug molecules.[79] Improved permeability across membranes has been observed for macrocyclic model compounds compared to acyclic ones. The improvement is assumed to be due to changes of properties, like reducing flexibility of macrocyclic compounds.[80] Shielding of polar groups by adjacent bulky lipophilic substituents has been designed to improve membrane permeability. Investigations showed that polar moieties, like amide bonds, can be shielded from the surrounding environment by lipophilic side chains to improve permeability.[81] The cyclic hexaleucine showed higher permeability and higher oral bioavailability in rat than the corresponding pentaleucine. The X-ray crystallography data demonstrated that the side chain in pentaleucine positioned outward from cyclic centre and exposed that the polar amide groups to hydrogen bonding with the solvent. On the contrary, the side chains of hexaleucine were towards the cycle center, shielding the amide protons and exhibiting cell permeability. In order to identify uncharged β-1,4-GalT inhibitors exhibiting good application in cell assays, strategies were investigated in our project for improving membrane permeability of inhibitors.

1.5. Project Objectives

β-1,4-GalTs are involved in many fundamental biological processes and β-1,4-GalT inhibitors are therefore sought after as tool compounds for chemical biology, drug discovery and biotechnology. Cellular applications of many existing β-1,4-GalTs inhibitors have been compromised by their limited stability and membrane permeability due to the presence of polar functional groups such as the charged pyrophosphate moiety. The main goal of this project was to develop uncharged β-1,4-GalT inhibitors with suitable properties for applications in cells.

In previous studies, 5-(5-formylthien-2-yl) UDP-galactose (5-FT UDP-Gal) has been reported as a potent β-1,4-GalT inhibitor, which acts by blocking the movement of a flexible loop in the active site during the catalytic cycle. The 5-substituent of 5-FT UDP-Gal is essential for this mode of action. In this project, several strategies were explored to develop novel, uncharged β-1,4-GalT inhibitors derived from 5-FT UDP-Gal (**Figure 26**).

In order to improve membrane penetration, a pro-drug concept using phosphate esters of 5-FT UDP-Gal was investigated. The aim was to mask the negatively charged pyrophosphate by conversion into esters that could be readily removed by intracellular carboxylesterases. In a related approach, nucleoside derivatives were developed, which contain an optimised 5-substituent, but lack the pyrophosphate and galactose moieties. It was reasoned that the membrane permeability of these nucleosides would be enhanced due to the absence of polar moieties, while the introduction of an optimised substituent at the 5-position could help to overcome the loss of binding affinity resulting from removal of the pyrophosphate and galactose moieties. These strategies are described in Chapter 2.

In order to overcome the potential loss of inhibitory activity in the nucleoside series resulting from the absence of the pyrophosphate and galactose moieties, a dynamic combinatorial

chemistry (DCC) strategy was investigated to identify suitable replacements of pyrophosphate and sugar moieties. Results from this approach are described in Chapter 3.

To facilitate the identification of active fragments in the DCC experiments, a positive control compound was required. A known acceptor substrate-based GalT inhibitor, GlcNAcβ-(2-naphthyl), was selected for these experiments. Unexpectedly, results from biochemical assays suggested that in contrast to published results, this compound behaved not an inhibitor but a substrate. This behaviour had not previously been reported, and a series of experiments was carried out to understand this discrepancy. These results are presented in Chapter 4.

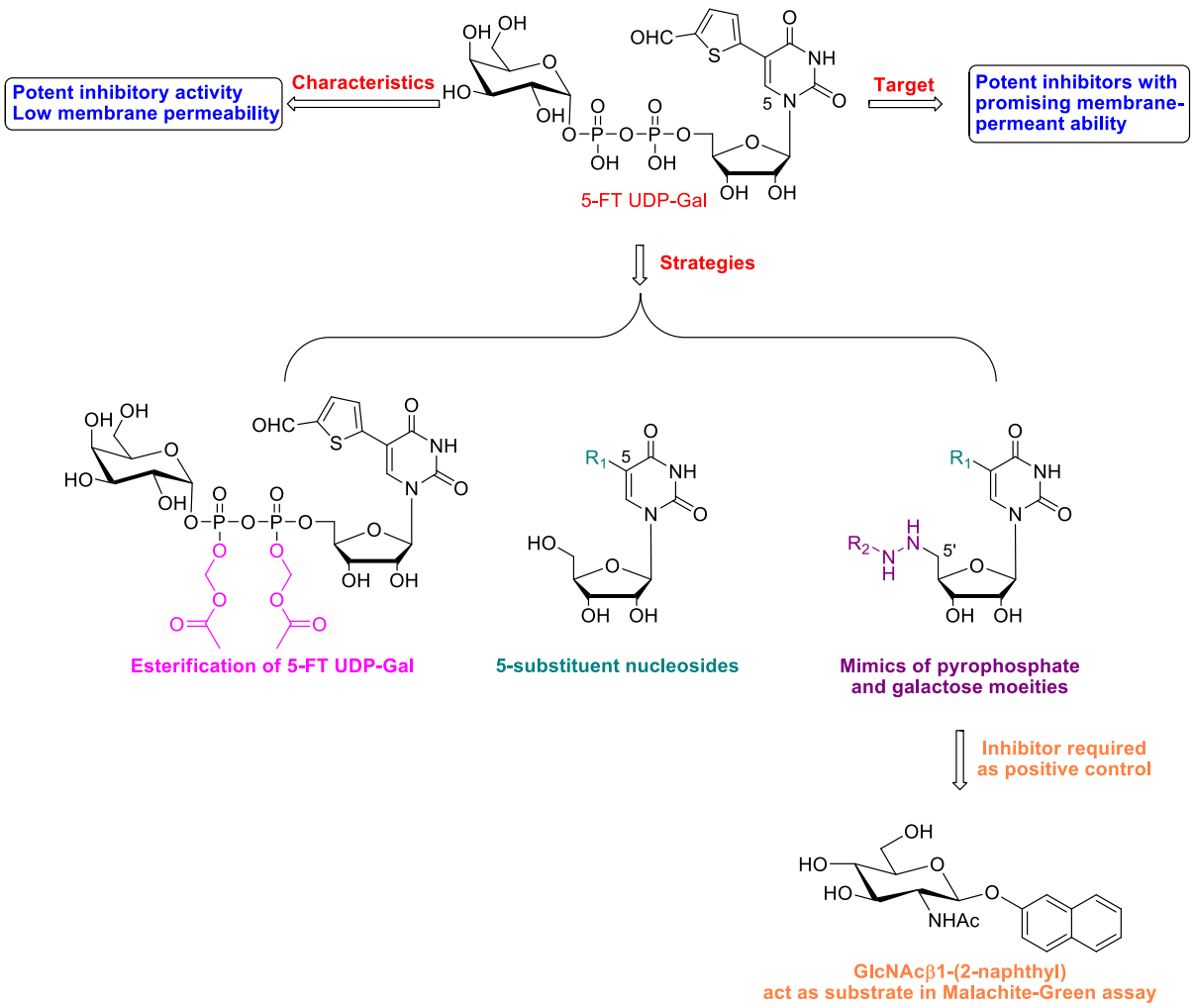


Figure 26 Illustration of project objectives: different strategies for the development of uncharged β-1,4-GalT inhibitors.

1.6. References

1. Breton, C.; Fournel-Gigleux, S.; Palcic, M. M., Recent structures, evolution and mechanisms of glycosyltransferases. *Curr Opin Struc Biol* **2012**, *22*, 540-549.

2. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G., Glycosyltransferases: Structures, functions, and mechanisms. *Annu Rev Biochem* **2008**, *77*, 521-555.

3. Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B., An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* **2003**, *328*, 307-17.

4. Lairson, L. L.; Chiu, C. P. C.; Ly, H. D.; He, S. M.; Wakarchuk, W. W.; Strynadka, N. C. J.; Withers, S. G., Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue. *J Biol Chem* **2004**, *279*, 28339-28344.

5. Persson, K.; Ly, H. D.; Dieckelmann, M.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C. J., Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs. *Nat Struct Biol* **2001**, *8*, 166-175.

6. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G., Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **2008**, *77*, 521-55.

7. Charnock, S. J.; Davies, G. J., Structure of the nucleotide-diphospho-sugar transferase, SpsA from Bacillus subtilis, in native and nucleotide-complexed forms. *Biochemistry* **1999**, *38*, 6380-6385.

8. Ramakrishnan, B.; Balaji, P. V.; Qasba, P. K., Crystal structure of beta 1,4-galactosyltransferase complex with UDP-Gal reveals an oligosaccharide acceptor binding site. *J Mol Biol* **2002**, *318*, 491-502.

9. Barrett, D.; Leimkuhler, C.; Chen, L.; Walker, D.; Kahne, D.; Walker, S., Kinetic characterization of the glycosyltransferase module of Staphylococcus aureus PBP2. *J Bacteriol* **2005**, *187*, 2215-2217.

10. Yuan, Y. Q.; Barrett, D.; Zhang, Y.; Kahne, D.; Sliz, P.; Walker, S., Crystal structure of a peptidoglycan glycosyltransferase suggests a model for processive glycan chain synthesis. *P Natl Acad Sci USA* **2007**, *104*, 5348-5353.

11. Igura, M.; Maita, N.; Kamishikiryo, J.; Yamada, M.; Obita, T.; Maenaka, K.; Kohda, D., Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *Embo Journal* **2008**, *27*, 234-243.

12. Leloir, L. F., Two decades of research on the biosynthesis of saccharides. *Science* **1971**, *172*, 1299-303.

13. Ernst, B.; Oehrlein, R., Substrate and Donor Specificity of Glycosyl Transferases. *Glycoconj J* **1999**, *16*, 161-170.

14. Baisch, G.; Öhrlein, R.; Katopodis, A., Enzymatic fucosylations of non-natural acceptors with non-natural donor-sugars. *Bioorg Med Chem Lett* **1997**, *7*, 2431-2434.

15. Baisch, G.; Öhrlein, R.; Katopodis, A., Enzymatic fucosylations with purine-diphosphate-fucoses (PDP-Fucoses). *Bioorg Med Chem Lett* **1996**, *6*, 2953-2956.

16. Yang, M.; Brazier, M.; Edwards, R.; Davis, B. G., High-throughput mass-spectrometry monitoring for multisubstrate enzymes: determining the kinetic parameters and catalytic activities of glycosyltransferases. *Chembiochem* **2005**, *6*, 346-57.

17. Khaled, A.; Ivannikova, T.; Auge, C., Synthesis of unnatural sugar nucleotides and their evaluation as donor substrates in glycosyltransferase-catalyzed reactions. *Carbohydr Res* **2004**, *339*, 2641-9.

18. Lairson, L. L.; Wakarchuk, W. W.; Withers, S. G., Alternative donor substrates for inverting and retaining glycosyltransferases. *Chem Commun* **2007**, 365-367.

19. Ramakrishnan, B.; Qasba, P. K., Crystal structure of lactose synthase reveals a large conformational change in its catalytic component, the beta1,4-galactosyltransferase-I. *J Mol Biol* **2001**, *310*, 205-18.

20. Ohtsubo, K.; Marth, J. D., Glycosylation in cellular mechanisms of health and disease. *Cell* **2006**, *126*, 855-67.

21. Rexach, J. E.; Clark, P. M.; Hsieh-Wilson, L. C., Chemical approaches to understanding O-GlcNAc glycosylation in the brain. *Nat Chem Biol* **2008**, *4*, 97-106.

22. Weerapana, E.; Imperiali, B., Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* **2006**, *16*, 91R-101R.

23. Abu-Qarn, M.; Eichler, J.; Sharon, N., Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Curr Opin Struc Biol* **2008**, *18*, 544-550.

24. Bertozzi, C. R.; Kiessling; L., L., Chemical Glycobiology. *Science* **2001**, *291*, 2357-2364.

25. Nothaft, H.; Szymanski, C. M., Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Micro* **2010**, *8*, 765-778.

26. Liu, H.; Ritter, T. K.; Sadamoto, R.; Sears, P. S.; Wu, M.; Wong, C. H., Acceptor specificity and inhibition of the bacterial cell-wall glycosyltransferase MurG. *Chembiochem* **2003**, *4*, 603-609.

27. Vollmer, W.; Blanot, D.; de Pedro, M. A., Peptidoglycan structure and architecture. *Fems Microbiol Rev* **2008**, *32*, 149-167.

28. Dube, D. H.; Bertozzi, C. R., Glycans in cancer and inflammation. Potential for therapeutics and diagnostics. *Nat Rev Drug Discov* **2005**, *4*, 477-488.

29. Orntoft, T. F.; Vestergaard, E. M., Clinical aspects of altered glycosylation of glycoproteins in cancer. *Electrophoresis* **1999**, *20*, 362-371.

30. Orntoft, T. F.; Greenwell, P.; Clausen, H.; Watkins, W. M., Regulation of the oncodevelopmental expression of type 1 chain ABH and Lewis(b) blood group antigens in human colon by alpha-2-L-fucosylation. *Gut* **1991**, *32*, 287-293.

31. Qasba, P. K.; Ramakrishnan, B.; Boeggeman, E., Structure and function of beta -1,4-galactosyltransferase. *Curr Drug Targets* **2008**, *9*, 292-309.

32. Ramakrishnan, B.; Boeggeman, E.; Qasba, P. K., Beta-1,4-galactosyltransferase and lactose synthase: molecular mechanical devices. *Biochem Biophys Res Commun* **2002**, *291*, 1113-8.

33. Ramakrishnan, B.; Boeggeman, E.; Ramasamy, V.; Qasba, P. K., Structure and catalytic cycle of beta-1,4-galactosyltransferase. *Curr Opin Struct Biol* **2004**, *14*, 593-600.

34. Qasba, P. K.; Ramakrishnan, B.; Boeggeman, E., Structure and Function of beta-1,4-Galactosyltransferase. *Curr Drug Targets* **2008**, *9*, 292-309.

35. Ramakrishnan, B.; Boeggeman, E.; Ramasamy, V.; Qasba, P. K., Structure and catalytic cycle of β-1,4-galactosyltransferase. *Curr Opin Struc Biol* **2004**, *14*, 593-600.

36. Ramasamy, V.; Ramakrishnan, B.; Boeggeman, E.; Qasba, P. K., The role of tryptophan 314 in the conformational changes of beta 1,4-galactosyltransferase-I. *J Mol Biol* **2003**, *331*, 1065-1076.

37. Gastinel, L. N.; Cambillau, C.; Bourne, Y., Crystal structures of the bovine β4galactosyltransferase catalytic domain and its complex with uridine diphosphogalactose. *The EMBO Journal* **1999**, *18*, 3546-3557.

38. Almeida, R.; Levery, S. B.; Mandel, U.; Kresse, H.; Schwientek, T.; Bennett, E. P.; Clausen, H., Cloning and Expression of a Proteoglycan UDP-Galactose: β-Xylose β1,4-Galactosyltransferase I: a seventh member of the human β4-galactosyltransferase gene family. *J Biol Chem* **1999**, *274*, 26165-26171.

39. Hennet, T., The galactosyltransferase family. *Cell mol life sciences : CMLS* **2002**, *59*, 1081-95.

40. Sperandio, M., Selectins and glycosyltransferases in leukocyte rolling in vivo. *Febs J*

2006, *273*, 4377-89.

41. Moloney, D. J.; Haltiwanger, R. S., The O-linked fucose glycosylation pathway: identification and characterization of a uridine diphosphoglucose: fucose-beta1,3-glucosyltransferase activity from Chinese hamster ovary cells. *Glycobiology* **1999**, *9*, 679-87.

42. Seko, A.; Dohmae, N.; Takio, K.; Yamashita, K., Beta 1,4-galactosyltransferase (beta 4GalT)-IV is specific for GlcNAc 6-O-sulfate. Beta 4GalT-IV acts on keratan sulfate-related glycans and a precursor glycan of 6-sulfosialyl-Lewis X. *J Biol Chem* **2003**, *278*, 9150-8.

43. Almeida, R.; Levery, S. B.; Mandel, U.; Kresse, H.; Schwientek, T.; Bennett, E. P.; Clausen, H., Cloning and expression of a proteoglycan UDP-galactose:beta-xylose beta1,4-galactosyltransferase I. A seventh member of the human beta4-galactosyltransferase gene family. *J Biol Chem* **1999**, *274*, 26165-71.

44. Hathaway, H.; Shur, B., Cell surface beta 1,4-galactosyltransferase functions during neural crest cell migration and neurulation in vivo. *J Cell Biol* **1992**, *117*, 369-382.

45. Han, Y.; Zhou, X.; Ji, Y.; Shen, A.; Sun, X.; Hu, Y.; Wu, Q.; Wang, X., Expression of beta-1,4-galactosyltransferase-I affects cellular adhesion in human peripheral blood CD4+ T cells. *Cell Immunol* **2010**, *262*, 11-17.

46. Rodeheffer, C.; Shur, B. D., Targeted mutations in β1,4-galactosyltransferase I reveal its multiple cellular functions. *Biochim et Biophy Acta (BBA) - General Subjects* **2002**, *1573*, 258-270.

47. Xu, C.; Liu, B.; Hu, B.; Han, Y.; Feng, L.; Allingham, J. S.; Szarek, W. A.; Wang, L.; Brockhausen, I., Biochemical Characterization of UDP-Gal:GlcNAc-Pyrophosphate-Lipid β-1,4-Galactosyltransferase WfeD, a New Enzyme from Shigella boydii Type 14 That Catalyzes the Second Step in O-Antigen Repeating-Unit Synthesis. *J Bacteriol* **2011**, *193*, 449-459.

48. Asano, M.; Furukawa, K.; Kido, M.; Matsumoto, S.; Umesaki, Y.; Kochibe, N.; Iwakura, Y., Growth retardation and early death of beta-1,4-galactosyltransferase knockout mice with augmented proliferation and abnormal differentiation of epithelial cells. *Embo J* **1997**, *16*, 1850-7.

49. Chen, L.; Xie, Y.; Fan, J.; Sui, L.; Xu, Y.; Zhang, N.; Ma, Y.; Li, Y.; Kong, Y., HCG induces β1,4-GalT I expression and promotes embryo implantation. *Int J Clin Exp Pathol* **2015**, *8*, 4673-4683.

50. Villegas-Comonfort, S.; Serna-Marquez, N.; Galindo-Hernandez, O.; Navarro-Tito, N.; Salazar, E. P., Arachidonic acid induces an increase of β-1,4-galactosyltransferase I expression in MDA-MB-231 breast cancer cells. *J Cell Biochem* **2012**, *113*, 3330-3341.

51. Asano, M.; Nakae, S.; Kotani, N.; Shirafuji, N.; Nambu, A.; Hashimoto, N.; Kawashima, H.; Hirose, M.; Miyasaka, M.; Takasaki, S.; Iwakura, Y., Impaired selectin-ligand biosynthesis and reduced inflammatory responses in β-1,4-galactosyltransferase-I–deficient mice. *Blood* **2003**, *102*, 1678-1685.

52. Yang, J.; Hirata, T.; Croce, K.; Merrill-Skoloff, G.; Tchernychev, B.; Williams, E.; Flaumenhaft, R.; Furie, B. C.; Furie, B., Targeted gene disruption demonstrates that P-selectin glycoprotein ligand 1 (PSGL-1) is required for P-selectin-mediated but not E-selectin-mediated neutrophil rolling and migration. *J Exp Med* **1999**, *190*, 1769-82.

53. Sperandio, M.; Smith, M. L.; Forlow, S. B.; Olson, T. S.; Xia, L.; McEver, R. P.; Ley, K., P-selectin glycoprotein ligand-1 mediates L-selectin-dependent leukocyte rolling in venules. *J Exp Med* **2003**, *197*, 1355-63.

54. Angiari, S.; Constantin, G., Regulation of T cell trafficking by the T cell immunoglobulin and mucin domain 1 glycoprotein. *Trends Mol Med* **2005**, *20*, 675-684.

55. Asano, M.; Nakae, S.; Kotani, N.; Shirafuji, N.; Nambu, A.; Hashimoto, N.; Kawashima, H.; Hirose, M.; Miyasaka, M.; Takasaki, S.; Iwakura, Y., Impaired selectin-ligand biosynthesis and reduced inflammatory responses in beta-1,4-galactosyltransferase-I-deficient mice. *Blood* **2003**, *102*, 1678-85.

56. Takaya, K.; Nagahori, N.; Kurogochi, M.; Furuike, T.; Miura, N.; Monde, K.; Lee, Y. C.; Nishimura, S.-I., Rational Design, Synthesis, and Characterization of Novel Inhibitors for Human β1,4-Galactosyltransferase. *J Med Chem* **2005**, *48*, 6054-6065.

57. Chung, S. J.; Takayama, S.; Wong, C.-H., Acceptor substrate-based selective inhibition of galactosyltransferases. *Bioorg Med Chem Lett* **1998**, *8*, 3359-3364.

58. Yuasa, H.; Palcic, M. M.; Hindsgaul, O., Synthesis of the carbocyclic analog of uridine 5′ -(α-D-galactopyranosyl diphosphate) (UDP-Gal) as an inhibitor of β(1 → 4)-galactosyltransferase. *Can J Chem* **1995**, *73*, 2190-2195.

59. Hayashi, T.; Murray, B. W.; Wang, R.; Wong, C. H., A chemoenzymatic synthesis of UDP-(2-deoxy-2-fluoro)-galactose and evaluation of its interaction with galactosyltransferase. *Bioorg Med Chem* **1997**, *5*, 497-500.

60. Takaya, K.; Nagahori, N.; Kurogochi, M.; Furuike, T.; Miura, N.; Monde, K.; Lee, Y. C.; Nishimura, S., Rational design, synthesis, and characterization of novel inhibitors for human beta1,4-galactosyltransferase. *J Med Chem* **2005**, *48*, 6054-65.

61. Wang, R.; Steensma, D. H.; Takaoka, Y.; Yun, J. W.; Kajimoto, T.; Wong, C. H., A search for pyrophosphate mimics for the development of substrates and inhibitors of

glycosyltransferases. *Bioorg Med Chem* **1997**, *5*, 661-72.

62. Murata, S.; Ichikawa, S.; Matsuda, A., Synthesis of galactose-linked uridine derivatives with simple linkers as potential galactosyltransferase inhibitors. *Tetrahedron* **2005**, *61*, 5837-5842.

63. Izumi, M.; Yuasa, H.; Hashimoto, H., Bisubstrate analogues as glycosyltransferase inhibitors. *Curr Top Med Chem* **2009**, *9*, 87-105.

64. Reck, F.; Springer, M.; Meinjohanns, E.; Paulsen, H.; Brockhausen, I.; Schachter, H., Synthetic substrate analogues for UDP-GlcNAc: Man alpha 1-3R beta 1-2-N-acetylglucosaminyltransferase I. Substrate specificity and inhibitors for the enzyme. *Glycoconj J* **1995**, *12*, 747-54.

65. Brown, J. R.; Yang, F.; Sinha, A.; Ramakrishnan, B.; Tor, Y.; Qasba, P. K.; Esko, J. D., Deoxygenated Disaccharide Analogs as Specific Inhibitors of β1–4-Galactosyltransferase 1 and Selectin-mediated Tumor Metastasis. *J Biol Chem* **2009**, *284*, 4952-4959.

66. Hashimoto, H.; Endo, T.; Kajihara, Y., Synthesis of the First Tricomponent Bisubstrate Analogue That Exhibits Potent Inhibition against GlcNAc:beta-1,4-Galactosyltransferase. *J Org Chem* **1997**, *62*, 1914-1915.

67. Chiang, C. H.; Wang, C. H.; Chang, H. C.; More, S. V.; Li, W. S.; Hung, W. C., A novel sialyltransferase inhibitor AL10 suppresses invasion and metastasis of lung cancer cells by inhibiting integrin-mediated signaling. *J Cell Physiol* **2010**, *223*, 492-9.

68. Matsson, P.; Doak, B. C.; Over, B.; Kihlberg, J., Cell permeability beyond the rule of 5. *Adv Drug Deliv Rev* **2016**, *101*, 42-61.

69. Ahlbach, C. L.; Lexa, K. W.; Bockus, A. T.; Chen, V.; Crews, P.; Jacobson, M. P.; Lokey, R. S., Beyond cyclosporine A: conformation-dependent passive membrane permeabilities of cyclic peptide natural products. *Future Med Chem* **2015**, *7*, 2121-30.

70. Wang, C. K.; Northfield, S. E.; Colless, B.; Chaousis, S.; Hamernig, I.; Lohman, R. J.; Nielsen, D. S.; Schroeder, C. I.; Liras, S.; Price, D. A.; Fairlie, D. P.; Craik, D. J., Rational design and synthesis of an orally bioavailable peptide guided by NMR amide temperature coefficients. *Proc Natl Acad Sci U S A* **2014**, *111*, 17504-9.

71. Wang, C. K.; Northfield, S. E.; Swedberg, J. E.; Colless, B.; Chaousis, S.; Price, D. A.; Liras, S.; Craik, D. J., Exploring experimental and computational markers of cyclic peptides: Charting islands of permeability. *Eur J Med Chem* **2015**, *97*, 202-213.

72. White, T. R.; Renzelman, C. M.; Rand, A. C.; Rezai, T.; McEwen, C. M.; Gelev, V. M.; Turner, R. A.; Linington, R. G.; Leung, S. S. F.; Kalgutkar, A. S.; Bauman, J. N.; Zhang, Y.; Liras, S.; Price, D. A.; Mathiowetz, A. M.; Jacobson, M. P.; Lokey, R. S., On-resin N-

methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nat Chem Biol* **2011**, *7*, 810-817.

73. Wang, P.; Chun, B. K.; Rachakonda, S.; Du, J.; Khan, N.; Shi, J.; Stec, W.; Cleary, D.; Ross, B. S.; Sofia, M. J., An efficient and diastereoselective synthesis of PSI-6130: a clinically efficacious inhibitor of HCV NS5B polymerase. *J Org Chem* **2009**, *74*, 6819-24.

74. Burnham, H. H., Some Difference in the action of ephedrine and ephedrine (adrenalin. etc.) on the nasal turbinate mucosa. *Can Med Assoc J* **1932**, *27*, 168-170.

75. Rezai, T.; Yu, B.; Millhauser, G. L.; Jacobson, M. P.; Lokey, R. S., Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J Am Chem Soc* **2006**, *128*, 2510-1.

76. Over, B.; McCarren, P.; Artursson, P.; Foley, M.; Giordanetto, F.; Gronberg, G.; Hilgendorf, C.; Lee, M. D. t.; Matsson, P.; Muncipinto, G.; Pellisson, M.; Perry, M. W.; Svensson, R.; Duvall, J. R.; Kihlberg, J., Impact of stereospecific intramolecular hydrogen bonding on cell permeability and physicochemical properties. *J Med Chem* **2014**, *57*, 2746-54.

77. Mallinson, J.; Collins, I., Macrocycles in new drug discovery. *Future Med Chem* **2012**, *4*, 1409-38.

78. Bogdan, A. R.; Davies, N. L.; James, K., Comparison of diffusion coefficients for matched pairs of macrocyclic and linear molecules over a drug-like molecular weight range. *Org Biomol Chem* **2011**, *9*, 7727-33.

79. Hill, T. A.; Lohman, R. J.; Hoang, H. N.; Nielsen, D. S.; Scully, C. C.; Kok, W. M.; Liu, L.; Lucke, A. J.; Stoermer, M. J.; Schroeder, C. I.; Chaousis, S.; Colless, B.; Bernhardt, P. V.; Edmonds, D. J.; Griffith, D. A.; Rotter, C. J.; Ruggeri, R. B.; Price, D. A.; Liras, S.; Craik, D. J.; Fairlie, D. P., Cyclic Penta- and Hexaleucine Peptides without N-Methylation Are Orally Absorbed. *ACS Med Chem Lett* **2014**, *5*, 1148-51.

CHAPTER 2

Uncharged nucleoside inhibitors of β-1,4-galactosyltransferase with activity in cells

2.1. Introduction

In this chapter, a brief introduction will be given to the pro-drug concept and Suzuki cross-coupling reactions. The latter was the main strategy for generating nucleoside-based GalT inhibitors. The main part of the chapter describes attempts to apply a pyrophosphate esterification strategy to 5-FT UDP-Gal, and the development of nucleoside derivatives of 5-FT UDP-Gal.

2.1.1. Ester pro-drugs of phosphate containing molecules

A pro-drug is an inactive precursor of an active compound, that can undergo a chemical or enzymatic conversion *in vivo* to release the active compound. Pro-drug approaches are applied to mask undesirable properties of active compounds, like low solubility in water or lipid membranes, low target selectivity, chemical instability, pre-systemic metabolism and toxicity.[1, 2] This is achieved by temporarily modifying the structure of the active drug with protecting groups and the prodrug is converted to the parent drug when it reaches the targets and the remnants of the protecting moieties are eliminated.

Phosphate esters are essential for all organisms.[3, 4] For example, Phosphates esters are central to information storage, serve as the currency for energy exchange, involve in intracellular signalling, and contribute to membrane fluidity. They are important intermediates in carbohydrate metabolism, in formation of nucleotides and their assembly into RNA and DNA, and in steroid fabrication and protein lipidation through the isoprenoid biosynthesis pathways. Due to its nature of negative charges, the phosphate ester containing compounds exhibit poor

membrane permeability, which prevent their intracellular application. Because of the high polarity, these compounds tend to be subject to efficient renal clearance and biliary excretion. Phosphates are substrates for many phosphatase, exhibiting enzymatic lability.[5] Therefore, effective prodrug approaches are developed towards phosphate containing compounds, which help to remove charge to enable them to transit biological membranes and then release as the active forms once inside the target cells.

To date, a wide variety of protecting groups for phosphates have been developed and they can be classified into two categories, esters[6, 7] and amides[8, 9]. Esterification is the most widely used prodrug design approach to mask charges or to increase the lipophilicity of poorly absorbed compounds. Ester prodrugs of phosphate are chemically stable and their synthesis are usually straightforward in the presence of a hydroxyl moiety. Besides the simplicity of chemical manipulation, the wide distribution as well as broad and overlapping substrate selectivity of esterases has contributed to the widespread use of phosphate ester prodrugs. (**Figure 27**)

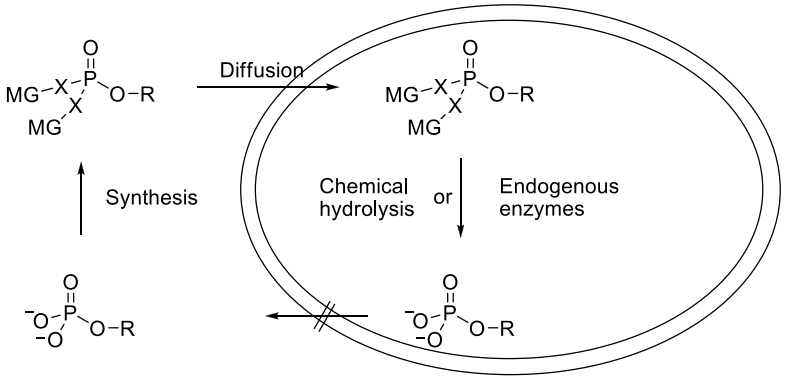


Figure 27 A suitable prodrug concept for phosphate esters: masking groups (MG) that allow diffusion across the plasma membrane are introduced by chemical synthesis. Inside cells the groups are removed and the phosphate ester may accumulate. (modified from reference[3])

The prodrug strategy was applied for the development of nucleotides for the treatment of HCV infection, the major cause of hepatocellular carcinoma.[10, 11] Some nucleoside derivatives were discovered to exhibit inhibition towards the HCV polymerase, which are essential enzyme for viral replication. The nucleosides are converted into the 5'-triphosphates to be active as inhibitors towards the enzyme, which undergo a series of phosphorylation steps catalysed by three kinases. (**Figure 28A**) However, the first phosphorylation of the antiviral nucleosides to the nucleoside 5'-monophosphates is the rate limiting step in human cells as the first kinase in the phosphorylation cascade is the most substrate selective. To avoid the problematic first phosphorylation step, the antiviral nucleosides usually are converted to its 5'-monophosphate form. The nucleoside monophosphate are poor drug candidates due to its poor membrane permeability. In addition, the monophosphates are susceptible to degradation by phosphatases. In order to deliver the antiviral nucleoside monophosphate, the ester prodrug strategies were carried out. For example, the bis(S-acyl-2-thioethyl) phosphate strategy (SATE) has been applied widely for the development of nucleoside 5'-phosphate prodrugs. The structure of nucleotide SATE prodrug consists of a nucleoside monophosphate where the phosphate group is masked by two S-acyl-2-thiothyl groups. The phosphate group is ultimately revealed by a sequence of enzymatic steps which requires esterase to cleave the thioesters and intramolecular attachment of the phosphate. (**Figure 28B**)

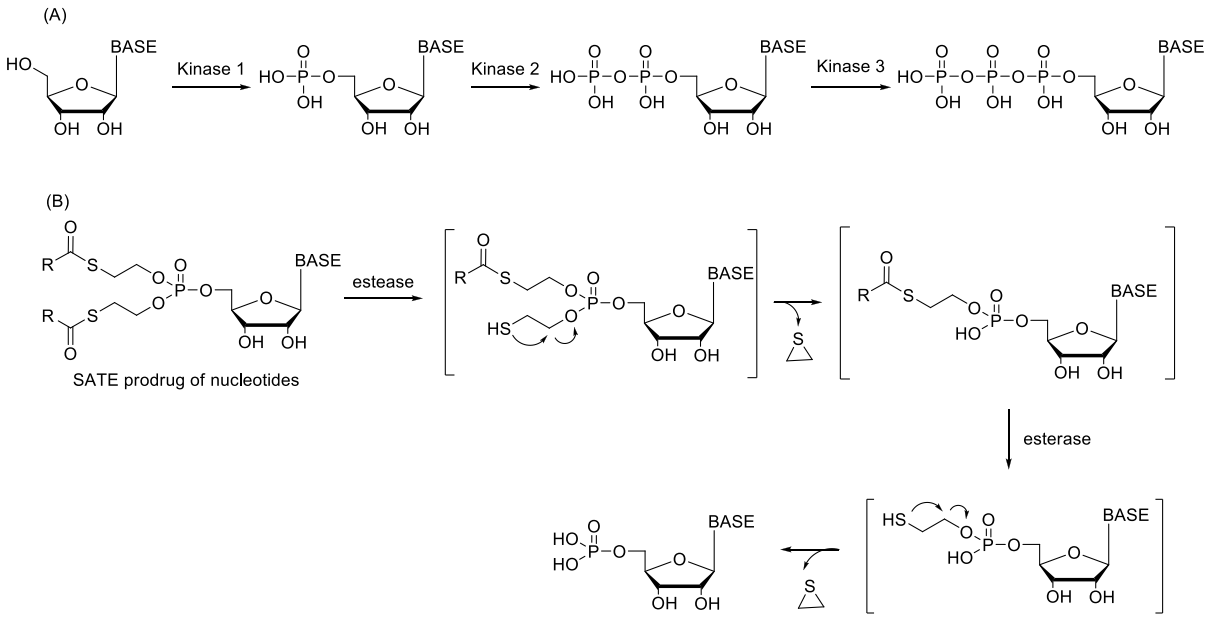


Figure 28 (A) Nucleoside kinase activation pathway resulting in nucleoside triphosphate which is the active substrate for a polymerase allowing incorporation of the nucleoside analogue into the growing RNA chain to exhibit inhibitory activity of virus replication. (B) The SATE prodrug of nucleotide decomposition pathway which produce the active compound nucleoside 5'-monophosphate.

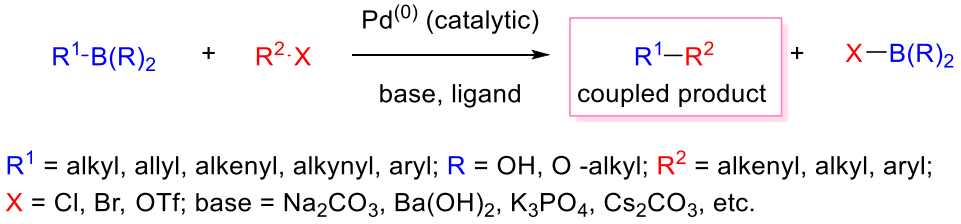
2.1.2. Synthesis of nucleoside derivatives *via* Suzuki cross-coupling

Nucleosides are compounds containing a nucleobase attached to a ribose or deoxyribose sugar ring by a glycosidic bond, such as adenosine, cytosine. Nucleoside derivatives are widespread in nature and various of the derivatives proves to exhibit different biological activities, like antiviral[12], antibacterial[13] and anticancer[14], and are explored as pharmaceutically active compounds.[15, 16] The applications of modified nucleosides are also universal. For example, 6-arylpurines are used as unnatural nucleobase in artificial base pairs.[17] C-8 modified nucleosides have shown potent antiviral and anticancer activity and have been studies as model compounds for DNA modification by carcinogenic compounds.[18] Fluorescently-labelled nucleosides are invaluable tools for the study of enzyme biological function.[19] Due to the broad biological activities and the innumerable application of

nucleosides, the development of a convenient synthetic protocol for the synthesis of nucleoside derivatives attracts increasing interest.

2.1.2.1. Suzuki cross-coupling chemistry

Cross-coupling reactions between organometallics and organic electrophiles, like halides, are versatile methods for carbon-carbon bond formation, such as Negashi cross-coupling[20], Heck reaction[21], Sonogashira cross-coupling[22]. Amongst these cross-coupling reactions, Suzuki-coupling reaction is one of the most widely characterized and developed reaction, which is carried out between organoboron reagents and organic electrophiles, like halides or triflates, mediated by palladium species under basic activation, forming carbon-carbon bond (**Scheme 1**).[23] The most commonly used organoboron reagents are boronic acids. The first example of Suzuki reaction was introduced by A. Suzuki and N. Miyaura.[24] Compared with other cross-coupling methodologies, Suzuki reaction offers several advantages, including the wide availability of organoboron and halide reagents, mild reaction conditions; compatibility with a wide variety of functional groups; water stability, less toxic *etc.*



Scheme 1 Typical Suzuki-Miyaura cross-coupling reaction

A general catalytic cycle for the Suzuki-Miyaura cross-coupling reaction encompasses four key steps, oxidative addition, ligand exchange, transmetallation and reductive elimination. The mechanism for the cross-coupling reaction is depicted in **Figure 29**.[25]

(a) The oxidative addition of an organic halide with the Palladium(0) species forms Palladium(II) species. The reactivity of palladium catalyst can be modulated by the ligands by donating or withdrawing electronic effect. The base present in the reaction serves a dual role in the catalytic cycle. The boron atom can be quarternised to form the reactive anion.

(b) The anion attached to the Pd(II) species is displaced by the anion of the base.

(c) The following step is the transmetallation between the Palladium species and the organoboron species. Pd(II) acts as an electrophile to react with nucleophiles and the organoboron speices are activated by the base to facilitate their nucleophilicity.

(d) Cross-coupling product is released from the palladium species via reductive elimination and the $Pd^{(0)}$ is regenerated as well for re-entering the catalytic cycle.

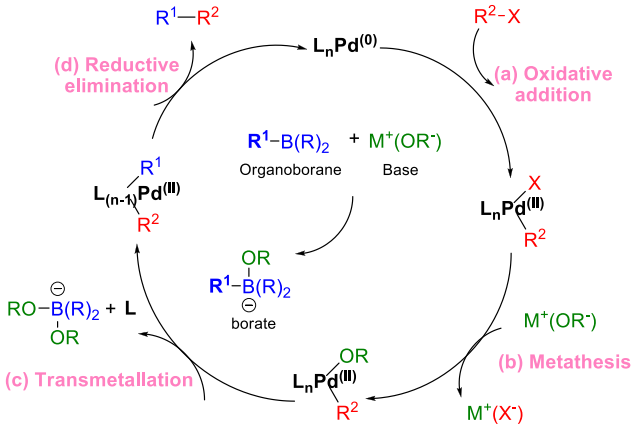


Figure 29 General catalytic cycle of the Suzuki-Miyaura cross-coupling reaction

The rate-determining step of Suzuki cross-coupling is usually believed to be the oxidative addition of organic halides towards palladium catalyst. The reactivity of compounds with different leaving groups decreases in the order Ar-I > Ar-Br > > Ar-Cl, which is determined by the stability of the leaving group that is being released. Moreover, other parameters of the cross-coupling also can influence the reaction, such as ligands, solvent, base, and palladium catalysts. Suzuki cross-coupling offers some disadvantages. For example, aryl halides generally react sluggishly; by-product such as self-coupling products are commonly formed; coupling products of phosphine-bound aryls are often formed.[26]

With the development of Suzuki cross-coupling reaction, it has risen in popularity for synthesis of natural products[27-29] and many libraries of compounds with biological activities[30-32]. As the inorganic-by-products are easily removed from the reaction mixture and the boronic acids are environmentally safer, the advantages have made the reaction scalable and cost-effective, suitable for industrial processes for drug synthesis.[33]

2.1.2.2. Synthesis of nucleoside derivatives *via* Suzuki cross-coupling

Suzuki cross-coupling chemistry is a versatile methodology applied to the modification of nucleobases in nucleoside derivatives.[34, 35] Suzuki reaction catalyses C-C bond formation between a boronic acid and an activated halide mediated by palladium species. Thus, in principle there are two strategies for the synthesis of nucleoside derivatives.[36] (**Figure 30**) The first one employs the cross-coupling between a halogenated nucleobase and an arylboronic acid. The second one involves the cross-coupling between a nucleoside boronic acid and a halogenoaryl species. Because nucleoside boronic acids are not commercially available and needed to be synthesized *via* the boronation of halogenated nucleosides, the

first strategy is more commonly carried out. Halogenated nucleoside derivatives can be prepared by standard electrophilic aromatic halogenation reactions.

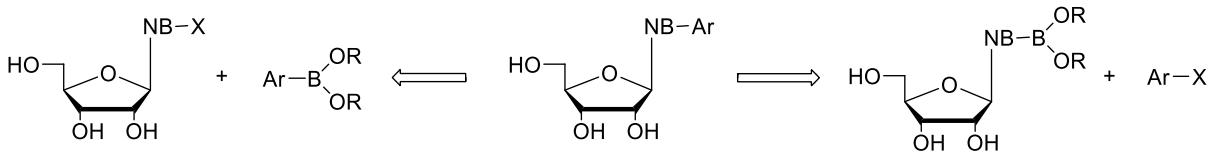


Figure 30 Two strategies of Suzuki cross-coupling for synthesis of nucleoside derivatives, NB = nucleobase

Although Suzuki cross-coupling reactions are well precedented for a wide range of substrates, their application to the synthesis of nucleoside derivatives still presents many challenges. When components containing functional groups which can coordinate to metal catalysts and deactive them are used in reaction, the yields are significantly reduced. A common approach is to protect the functional groups to prevent the coordination. Yet the introduction of extra two steps can result in low yield and poor atom economy.[37] The polar nature of nucleoside derivatives usually result in poor solubility in typical organic solvents. In order to solve these problems that lead to ineffective synthesis of nucleoside analogues, an approach by utilizing unprotected nucleosides for Suzuki reaction in aqueous media was developed.[34, 35] The aqueous solution, only water or water/organic co-solvent, can dissolve components well, leading to increasing conversion to desired products. Also, the direct structural modification of unprotected nucleosides affords an efficient approach.[38-40]

The electronic characteristics of components used for synthesis of nucleoside derivatives influence the cross-coupling. For example, electron-rich halogenated heterocycles, like furans, pyrroles, undergo slow oxidative addition to metal complexes. The boronic acid containing electron-deficient groups exhibit poor nucleophilic ability toward Pd species in the

transmetallation step. New ligands and solvent/base systems have been developed to facilitate the cross-coupling reaction.[34]

The microwave-assisted Suzuki cross-coupling approach was also applied to the synthesis of nucleoside analogues. Microwave-assisted synthesis has attracted much interest because of the possibility of drastically reducing the reaction time and improvement of reaction yields in comparison to conventional reaction methods. The application of microwave heating to palladium catalysed Suzuki reactions exert a similarly accelerating effect on the catalytic reaction, resulting in a completion in a short time. It has been shown that microwave irradiation can significantly promote the Suzuki-miyaura cross-coupling reaction. For example, Suzuki cross-coupling was applied for the substituted 2'-deoxyuridine and the reaction involved unprotected 5-iodo-2'-deoxyuridine and boronic acids.[41] The best conditions was utilized in the presence of Na_2CO_3 (1 equiv.), $Pd(OAc)_2$ (3 mol %), and PPh_3 (5.4 mol%) in water at 80 °C for several hour. Microwave irradiation at 120 °C for 10 min was applied in the presence of the reagents described above. The change of heating mode affords a significant reduce in the reaction time (from hours to 10 min) while keeping the same efficiency. (**Figure 31**)

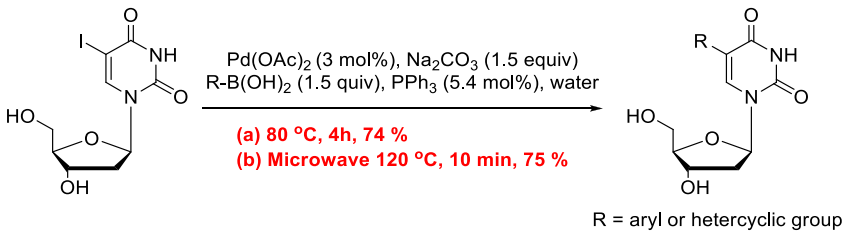


Figure 31 Suzuki cross-coupling for synthesis of substituted 2'-deoxyuridine via conventional heating or microwave condition.

2.2. Objectives:

A new class of GalT inhibitors, UDP-Gal derivatives with an additional substituent at uracil base, has been identified recently.[42, 43] The prototype inhibitor of this series, 5-formylthien-2-yl UDP-Gal (5-FT UDP-Gal, **Figure 32**) exhibits potent inhibitory activity towards several GalTs with K_i values in the low micromolar range. It is suggested that the 5-position substituent is essential for inhibitory activity, since it interferes with the movement of a flexible loop of enzyme active site from the open to the closed conformation.[43, 44] This closed conformation is required for the full catalytic activity during the catalytic cycle, therefore, 5-FT UDP-Gal is a promising inhibitor against GalTs. However, the pyrophosphate linkage is unstable and its membrane permeability is limited so that the application of 5-FT UDP-Gal is compromised due to the inherent negative charge of the pyrophosphate linkage. Several strategies were explored based on 5-FT UDP-Gal in order to develop β-1,4-GalT inhibitors with promising activity in cells by facilitating membrane permeability. (**Figure 32**)

A pro-drug concept was investigated in the first approach using a pyrophosphate ester of 5-FT UDP-Gal. Its membrane permeability could increase remarkably if the negative charge of sugar nucleotide was removed and maintaining the pyrophosphate moiety, which was essential for binding the enzymes, would facilitate potent binding affinity between enzymes and compounds. Then, the ester group could be readily removed by the related intracellular enzyme. The second approach entailed developing nucleoside-based derivatives as β-1,4-GalTs inhibitors by removing pyrophosphate and sugar moieties and introducing substituents in the 5-position of uracil. The membrane permeability of derivatives could be facilitated due to the absence of polar moieties. It was speculated that the introduction of the 5-position

substituent could play an important role for binding with the enzymes to overcome the loss of binding affinity from the pyrophosphate and galactose moieties.

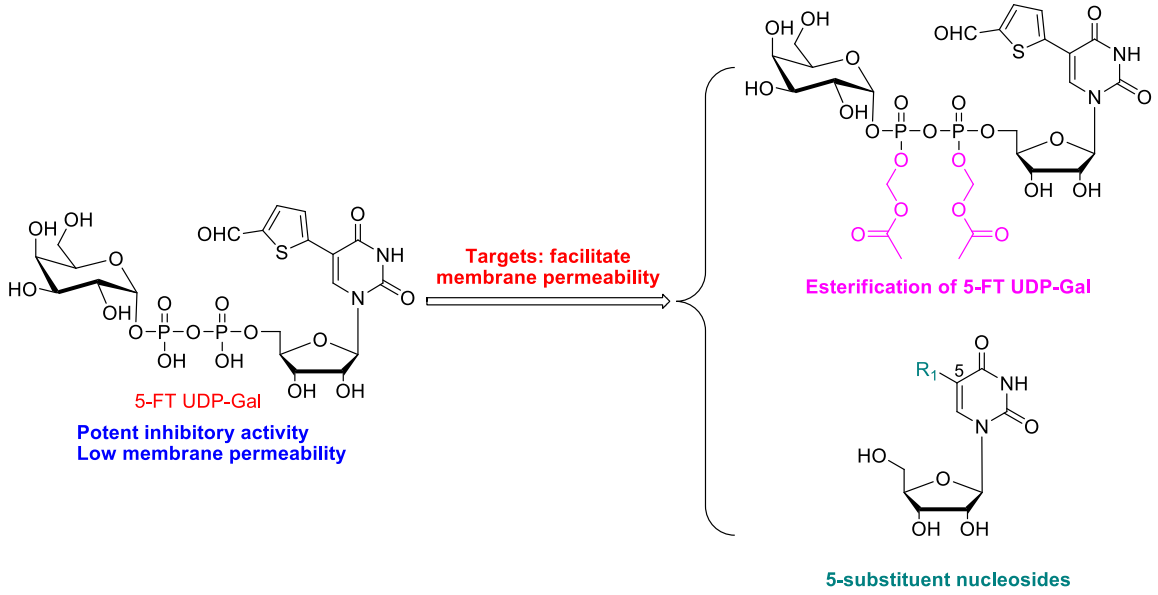
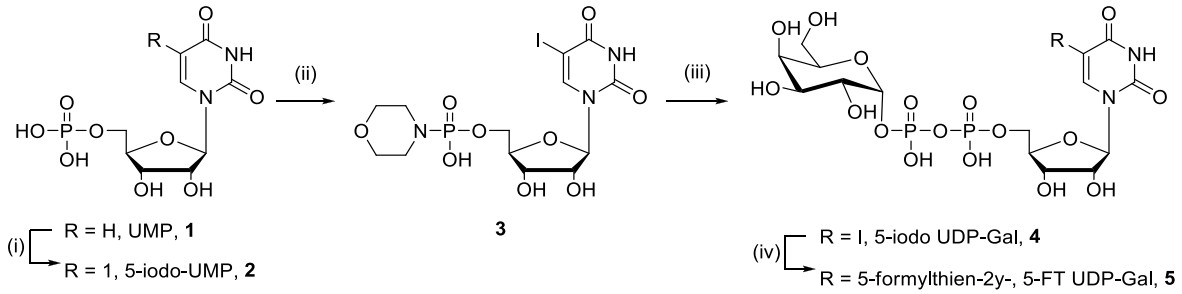


Figure 32 Design of potent inhibitors with promising membrane permeability towards β-1,4-GalTs

2.3. Esterification of sugar-nucleotides

In order to start the esterifcation of 5-FT UDP-Gal, the synthesis of 5-FT UDP-Gal was carried out firstly. (**Scheme 2**) Initiating from the iodination of uridine monophosphate (UMP) **1** under a harsh reaction condition by using 2M HNO_3 at 80 ºC for 12h, 5-iodo uridine monophosphate (UMP) **2** was obtained with a yield of 75 %. The subsequent phosphormopholidation was performed under the Mukaiyama condition. After stirred at room temperature for 1.5h, the desired phosphoro mopholidate **3** was precipitated as a white powder by the addition of a solution of NaI in acetone. Then compound **3** was coupled with galactose-α-phosphate with the catalyst of N-methylimidazole hydrochloride (NMICl),

forming 5-iodo UDP-Gal **4** in 48 % yield. The 5-formylthien-2-yl substituent was introduced by Suzuki coupling reaction, with an isolated yield of 45 %.

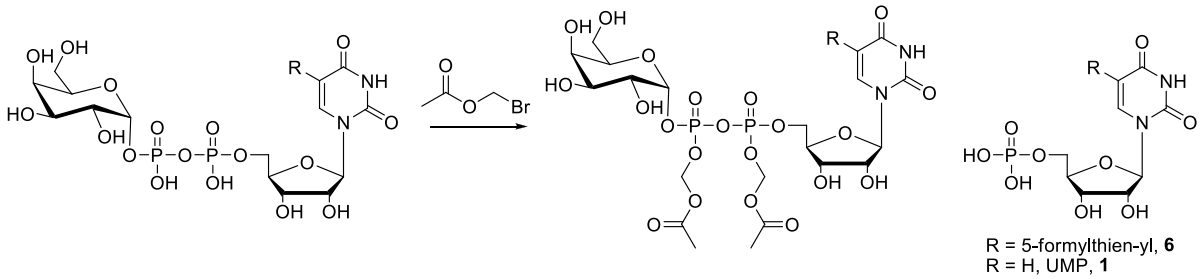


Scheme 2 Synthesis of 5-FT UDP-Gal. Conditions and reagents: (i) I_2, HNO_3 (2 M), $CHCl_3$, 80 °C, 12h, 75 %. (ii) Morpholine, 2,2'-dipyridyldisulfide, PPh_3, DMSO, rt, 1.5h, 67 %. (iii). Gal-α-monophosphate, N-methylimidazole hydrochloride (NMICl), MeCN, DMF, rt, 9h, 48 %. (iv) Boronic acid, Cs_2CO_3, Tris(3-sulfophenyl)phosphine trisodium salt (TPPTs), $PdNa_2Cl_4$, degassed water, 50 °C, 1.5h, 45 %.

With compound **5** in hand, the esterified reaction was employed. (**Table 1**) Initial attempt for esterification of sugar nucleotide was tested under the condition by using 20 equiv. N, N-Diisopropylethylamine (DIPEA), 5 equiv. bromo-methylacetate in anhydrous DMF solution at rt for 1d. (**entry 1**) TLC monitored that the starting material consumed and new spots appeared. (Mobile phase: ammonia/isopropanol/water = 1/3/6) After reverse phase chromatography purification, only pyrophosphate cleavage product **6** was obtained and sugar phosphate part was not collected from the eluent due to its UV inactivity. The obtained product **6** supported the decomposition of 5-FT UDP-Gal. Considering that the basic condition may interfere with the stability of pyrophosphate moiety, less excess equivalent of DIPEA was utilized in the reaction. (**entry 2**) When 5 equiv. base was utilized in the esterification, the starting material was still decomposed. It was speculated that the pyrophosphate could not survive for such a long reaction time due to its instability. Thus, the esterification of 5-FT UDP-Gal was carried out at rt for only 1h. The use of shorter reaction time also did not result in any product as it was shown on the TLC that the decomposition

took place but no esterified product was formed. (**entry 3**) The use of anhydrous MeCN as the replacement solvent of DMF also did not result in any product. (**entry 4**).

Table 1 Attempts for esterification of 5-FT UDP-Gal



Entry	R	Solvent	Base	Time	Result
1	5-formylthien-2-yl	Dry DMF	20 equiv DIPEA	1d	Decomposition
2	5-formylthien-2-yl	Dry DMF	5 equiv DIPEA	1d	Decomposition
3	5-formylthien-2-yl	Dry DMF	5 equiv DIPEA	1h	Decomposition
4	5-formylthien-2-yl	Dry MeCN	5 equiv DIPEA	1h	Decomposition
5	5-formylthien-2-yl	Dry DMF	5 equiv TEA	1h	Decomposition
6	I	Dry DMF	5 equiv DIPEA	1h	Decomposition
7	H	Dry DMF	5 equiv DIPEA	1h	Decomposition

In order to investigate the effect of base towards esterification, a different base, triethylamine (TEA), was employed for the reaction. The condition used to carry out the esterification involved 5 equiv. bromo methylacetate, 5 equiv. TEA as the base in DMF for 1h, as shown in **entry 5**. Under this condition, the formation of desired product similarly did not succeed. It was speculated that the 5-subtituent of uracil moiety might influence the reactivity of 5-FT UDP-Gal. To verify it, the intermediate, 5-iodo UDP-Gal was then utilized, whereas this attempt under condition of 5 equiv. DIPEA as base in DMF solution for 1h still failed. (**entry 6**) The unsubstituted compound, commercial UDP-Gal was used as starting material. (**entry 7**)

Before utilized in the reaction, the commercial material was pre-treated *via* an ion-exchange column to exchange the sodium ion into triethylamine salt. It is thought that the strong electrostatic interaction between sodium ion and phosphate bond might interfere with the esterification. However, only decomposed product UMP was obtained.

As the starting materials used in the reaction were all triethylamine salt, which were obtained *via* reverse phase chromatography with triethylammonium bicarbonate buffer (TEAB) and methanol as mobile phase, it was hypothesised that the interaction between triethylammonium cation and phosphate still prevent the formation of esterified product. The compound **7** containing pyrophosphate free acid would exhibit better reactivity. Therefore, the commercial UDP-Gal material, disodium UDP-Gal, was applied as a model reaction. Dowex 50WX2 hydrogen form resin was utilized for proton exchanging, leading to free acid of UDP-Gal. (**Figure 33**) Before utilised for ion exchange, the resin was pre-treated. This commercial resin was washed by methanol and deionized water separately for the removal of impurities. Then 1 M HCl solution was used to wash and saturate the resin, which was monitored by pH paper. The excess acid was removed by washing with deionized water. Disodium UDP-Gal was employed for proton exchange with the pre-treated resin and the free acid material was collected. With **7** in hand, it was used for the esterification under the condition of 5 equiv. DIPEA in anhydrous DMF solution for 1h. To our disappointment, the reaction still failed.

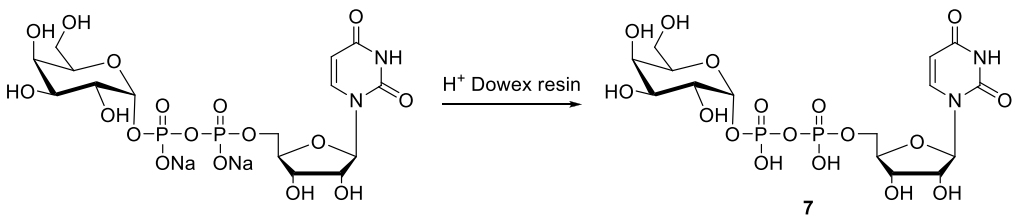


Figure 33 Proton exchange of UDP-Gal by cation Dowex resin

It was reported that the pyrophosphate linkage of UDP-GlcNAc could be mono-esterified.[45] The difference of structures between UDP-GlcNAc and UDP-Gal is the 2-position substituent of galactose. It was proposed that the unprotected hydroxyl group was more active comparing with acetylated NH group. It could take place the nucleophilic attack toward pyrophosphate bond, facilitating the cleavage of pyrophosphate moiety. (**Figure 34**) Due to the high reactivity of 2'-OH group, it was difficult to carry out the esterification of galactose-nucleotides directly. Because of these practical difficulties in obtaining the desired phosphate esters, it was decided to turn our attention to alternative strategies for the development of membrane-permeant inhibitors.

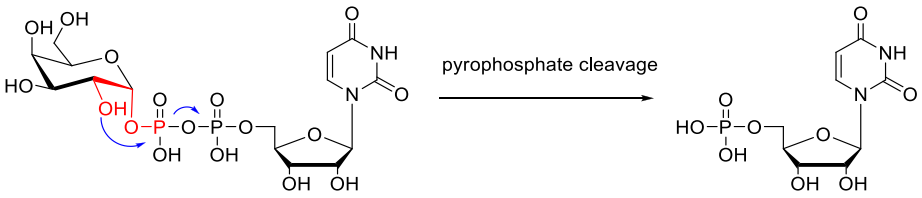


Figure 34 Proposed mechanism of pyrophosphate cleavage

2.4. Synthesis of nucleoside derivatives:

It is known that the substituent in 5-position is crucial for GalT inhibitory activity as it blocks the movement of flexible loop in the enzyme active site.[46] Therefore, in our initial attempt to obtain an uncharged inhibitor of β-1,4-GalTs, we retained the 5-substituted uridine fragment of 5-FT-UDP-Gal, and removed the pyrophosphate and sugar moieties (**Figure** 35).

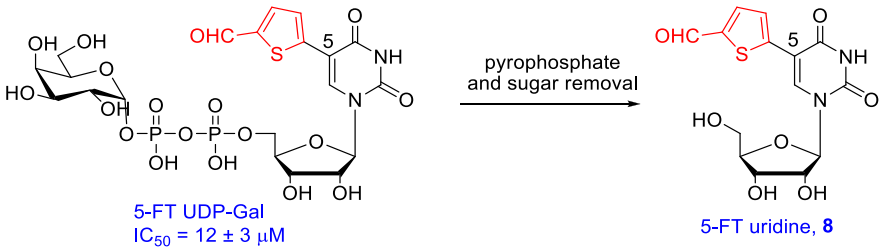
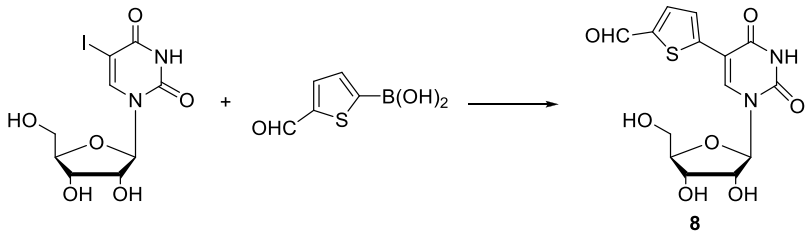


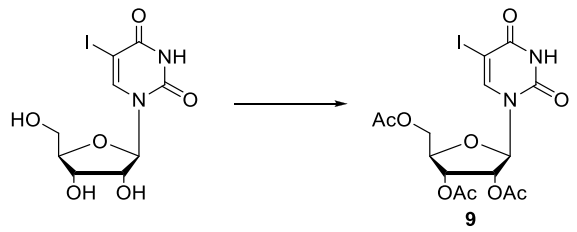
Figure 35 Design of 5-FT uridine modelled from the structure of 5-FT UDP-Gal

For the synthesis of the initial target compound 5-formylthien-2-yl uridine (5-FT uridine, **8**), previously established Suzuki cross-coupling conditions using conventional heating were employed. (**Scheme 3**) 5-iodo uridine was used as starting material. Despite a reaction time of up to 2 days, **8** was obtained in only 26% yield.



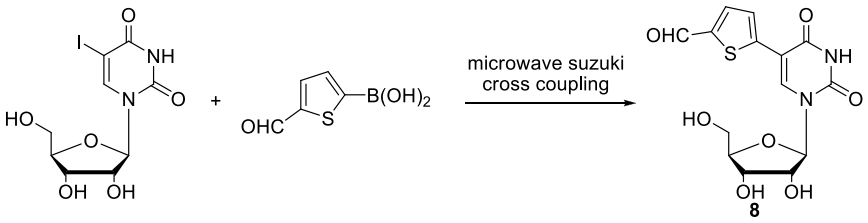
Scheme 3 Synthesis of 5-FT uridine. Reagents and conditions: boronic acid, Cs_2CO_3, $PdNa_2Cl_4$, TPPTs, degassed water, 2d, rt, 26 %.

Considering that the free hydroxyl groups of unprotected nucleoside may interfere with the catalyst reactivity, the 5-iodo uridine was acetylated prior to utilized in Suzuki reaction (**Scheme 4**). When acetylated 5-iodo uridine **9** was utilized in the Suzuki cross-coupling, the yield was even lower. It was speculated that the poor water solubility of starting material lead to the drop of yield.



Scheme 4 Synthesis of acetylated 5-iodo uridine. Reagents and conditions: Ac_2O, triethylamine, 4-dimethylaminopyridine, MeCN, rt, 3h, 89%.

It is known that due to the rapid heating technology of microwave irradiation, the reaction time for synthesis of desired product could be reduced drastically to a few minutes and the palladium catalysed processes could be facilitated compared with conventional heating. Therefore, the microwave Suzuki cross-coupling was carried out for the synthesis of 5-substitute nucleoside analogues (**Scheme 5**). For 5-formylthien-2-yl boronic acid, microwave condition led to markedly improved yield, from 26 % to 54 %, and significantly reduced reaction time, from 2 days to 0.5h, compared to conventional heating.



Scheme 5 Synthesis of 5-FT uridine. Reagents and condition: boronic acid, Cs_2CO_3, $PdNa_2Cl_4$, TPPTs, degassed water, MW 120 ^oC, 30min, 54%.

With this product **8** in hand, its inhibitory activity was assessed in a phosphatase-coupled glycosyltransferase assay (for details of this assay see Chapter 4). However, while the 5-FT UDP-Gal inhibits β-1,4-GalTs with an IC_{50} value of around 12 μM, the corresponding nucleoside derivative **8** was inactive at concentration up to 1 mM. In order to overcome this

dramatic drop in inhibitory activity resulting from removal of the pyrophosphate linkage and sugar moieties, the optimization of the substituent in the 5-position was carried out. A series of uridine derivatives with different heteroaromatic substituent on position 5 was designed and synthesis (**Figure 36**).

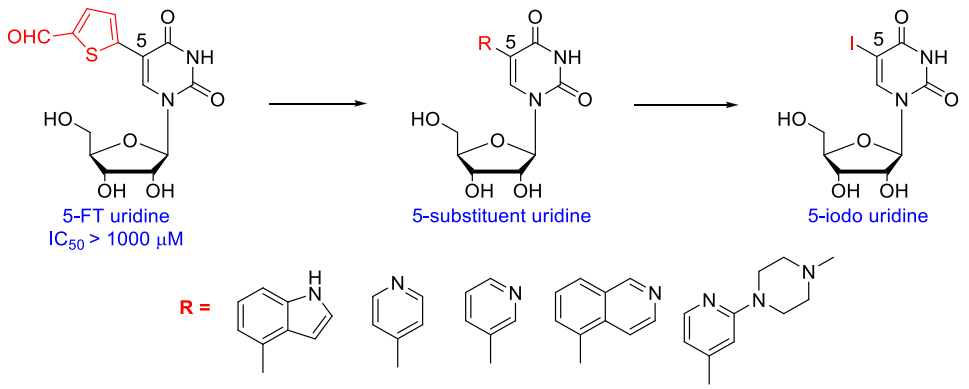


Figure 36 Uridine derivatives with substituents in 5-position

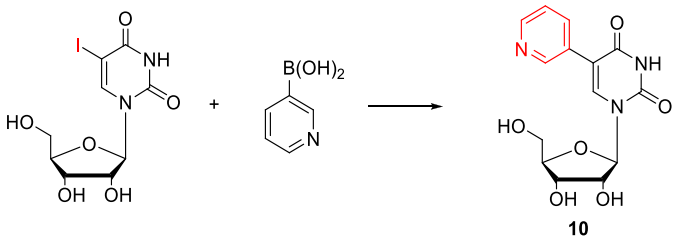
According to their structures, these target molecules also can be obtained by Suzuki-Miyaura cross-coupling of 5-iodo uridine with different aryl and heteroaryl boronic acids. By analysis of the Suzuki coupling mechanism, in the transmetallation step, $Pd^{(II)}$ acts as electrophiles to react with nucleophiles. (**Figure 29**) As the organoboron reagent tends to be the nucleophilic component in the catalytic cycle, electron-deficient (hetero)aryl boronic acids will react sluggishly under traditional conditions. In order to facilitate the cross-coupling of such less reactive boronic acids, microwave conditions for Suzuki-miyaura conditions between 5-iodo uridine and boronic acids were explored.

Initial trial of the Suzuki cross-coupling reaction for synthesis of 5-pyridinyl-3-yl uridine **10** was tested under the same experimental condition used for **8**. The reaction was carried out by using boronic acid (1.5 equiv.), Cs_2CO_3 (2 equiv.), TPPTs (0.0625 equiv.), $PdNa_2Cl_4$ (0.025

equiv.) in degassed water at MW 120 ^{o}C (**Table 2, entry 1**). After 30min, only a mixture of starting material and the deiodinated compound obtained. The use of microwave irradiation for longer time (**entry 2**) also did not result in any product. The formation of deiodinated product suggested that the oxidative addition step of Suzuki reaction took place, yet due to the poor nucleophilicity of electron-deficient pyridine-3-boronic acid, the transmetalation between boronic acid and palladium species failed. We then investigated the Suzuki cross-coupling reaction condition with different palladium species. The condition used to carry out the cross-coupling reaction involved tetrakis (triphenylphosphine palladium$^{(0)}$) (5 mol%) as catalyst, potassium carbonate (3 equiv.) as the base, and boronic acid (1.5 equiv.) in dimethyl ether (DME)-ethanol (9:1), as shown in **entry 3**. Under this condition, the formation of desired product similarly did not success. The use of microwave irradiation at 150 ^{o}C and toluene-ethanol (2:1) as solvent also did not result in any product. (**entry 4**) It is known that cesium base is generally effective base for Suzuki cross-coupling.[47] According to the mechanism, in the case of strong base with large size cation, the nucleophilicity of the boronate anion will be increased and more reactive towards the palladium species, which can accelerate the rate of transmetalation step. To verify this effect, pyridine-3-boronic acid underwent the Suzuki coupling in the presence of Cs_2CO_3 (2 equiv.), teterakis (5 mol%), 1,2-dimethoxyethane-water (5;1) as solvent, whereas no desired product was obtained (**entry 5**). It is known that the addition of fluoride ion could facilitate the cross-coupling process *via* the possible formation of a stable fluoroborate intermediate in the transmetalation step.[48] However, attempts to cross-couple pyridine-3-boronic acid under conditions, $Pd(OAc)_2$, KF (2 equiv.) and MeOH as solvent, completely failed (**entry 6**). Reaction was carried out at MW 130 ^{o}C for 30 min using a mixture of pyridine-3-boronic acid (1.5 equiv.), $PdCl_2(dppf)CH_2Cl_2$ (5 mol%), $NaHCO_3$(3 equiv.), DME/H_2O (3;1) (**entry 7**). A 20% yield to the cross-coupling product was observed. It was speculated that in the palladium species, the

1,1’-Bis(diphenylphosphino)ferrocene (dppf) ligand can enhance the reactivity towards boronate intermediate and dramatically improve the efficiency of cross-coupling reaction.

Table 2 Suzuki-Miyaura coupling of 5-iodo uridine with pyridine-3-boronic acid



10

Entry	Reagents	Microwave	Results
1	$PdNa_2Cl_4$, TPPTs, Cs_2CO_3, H_2O	MW 120 ºC, 0.5h	deiodination
2	$PdNa_2Cl_4$, TPPTs, Cs_2CO_3, H_2O	MW 120 ºC, 1h	deiodination
3	$Pd(PPh_3)_4$, K_2CO_3, DME/EtOH = 9/1	MW 120 ºC, 0.5h	deiodination
4	$Pd(PPh_3)_4$, K_2CO_3, Toluene/EtOH = 2/1	MW 150 ºC, 0.5h	deiodination
5	$Pd(PPh_3)_4$, Cs_2CO_3, DME/H_2O = 5/1	MW 120 ºC, 1h	deiodination
6	$Pd(OAc)_2$, KF, MeOH	MW 120 ºC, 0.5h	deiodination
7	$PdCl_2$(dppf)DCM, $NaHCO_3$, DME/H_2O = 3/1	MW 130 ºC, 0.5h	20 %

After obtaining compound **10** successfully, a second derivative **11**, 5-pyridinyl-4-yl uridine, was prepared in the same way starting from 5-iodo uridine and pyridine-4-boronic acid (1.5 equiv.). After purification of the crude product, the compound **11** was obtained in 27 % yield (**Table 3**, **entry 1**). Since the difference between compound **14** and **10** is only the additional substituent in the 2’-position of the pyridine ring, the Suzuki cross-coupling was performed

with 2-(4-methylpiperazin-1-yl)-pyridine-4-boronic acid (1.5 equiv.) using the same previously described condition ($PdCl_2(dppf)CH_2Cl_2$ (5 mol%), $NaHCO_3$ (3 equiv.), DME/H_2O (3;1), MW 130 ^oC, 30 min. The analogue **14** was obtained in 35 % yield (**entry 4**). Then the synthesis of compound **12** and **13** were started from 5-iodo uridine and isoquinoline-4-boronic acid and indole-4-boronic acid separately using TPPTs (0.0625 equiv.), $PdNa_2Cl_4$ (0.025 equiv.) as ligand and catalyst, Cs_2CO_3 (2 equiv.) as the base in degassed water at MW 120 ^oC for 30 min. The derivatives **12** and **13** were obtained in 25% and 43% yield separately (**entry 2** and **3**)

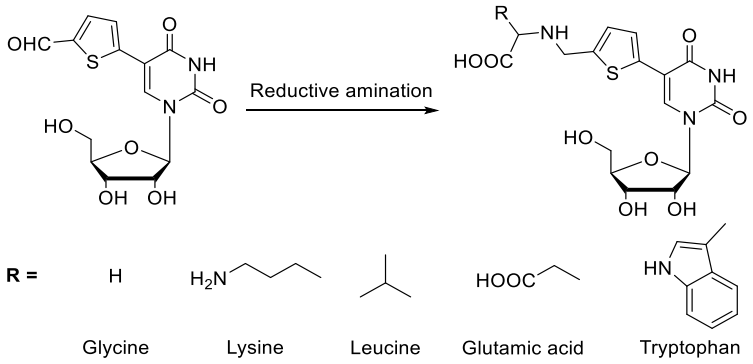
Table 3 Suzuki-Miyaura coupling of 5-iodo uridine with different heteroarylbornic acids R-$B(OH)_2$

Entry	R	Cmpd	Reagents	Microwave	Yield (%)
1	5-pyridinyl-4-yl-	**11**	$PdCl_2(dppf)DCM$, $NaHCO_3$, DME/H_2O (3:1)	MW 130 ^oC, 0.5h	27
2	5-isoquinoline-4-yl-	**12**	$PdNa_2Cl_4$, TPPTs, Cs_2CO_3, H_2O	MW 120 ^oC, 0.5h	25
3	5-indole-4-yl-	**13**	$PdNa_2Cl_4$, TPPTs, Cs_2CO_3, H_2O	MW 120 ^oC, 0.5h	43
4	5-(2-(4-methylpiperazin-1-yl)-pyridine-4-yl-	**14**	$PdCl_2(dppf)DCM$, $NaHCO_3$, DME/H_2O (3:1)	MW 130 ^oC, 0.5h	35

2.5. Synthesis of amino derivatives of 5-FT uridine via reductive amination

In order to generate additional structure diversity, compound **8** was used for reductive amination with a range of amino acids (**Scheme 6**). It was speculated that through the extension of the 5-substituent, these derivatives might be able to engage in additional interactions with the target enzyme β-1,4-GalTs as more functional groups would be included.

For the modifications, natural amino acids with different side chain functionalities were chosen, including L-glutamic acid, L-lysine, L-tryptophan, L-glycine and L-valine.



Scheme 6 Synthesis of reductive aminated derivatives. Reagents and conditions: amino acids, $NaBH_3CN$, MeOH, rt, 24h.

The initial trial of reductive aminated reaction was tested between **8** and Lysine (3 equiv.). The reaction was carried out by using $NaBH_3CN$ (5 mol %) in methanol at room temperature. After 1d, a mixture of the reductive aminated product **15** was obtained with a yield of 68 % (**Table 4**, **entry 1**) because there are two free amino groups in lysine. According to the ^{1}H-NMR spectra (**Figure 37**), the ratio between these two compounds is 5:1. Since these two compounds are structurally similar, their R_f values on TLC plate are also similar and it is difficult to separate them through flash chromatography. Separation was ultimately achieved by semi-preparative HPLC, and the compounds were identified unambiguously through the heteronuclear correlations between C_a and H_1/H_2 in their HMBC NMR spectra.

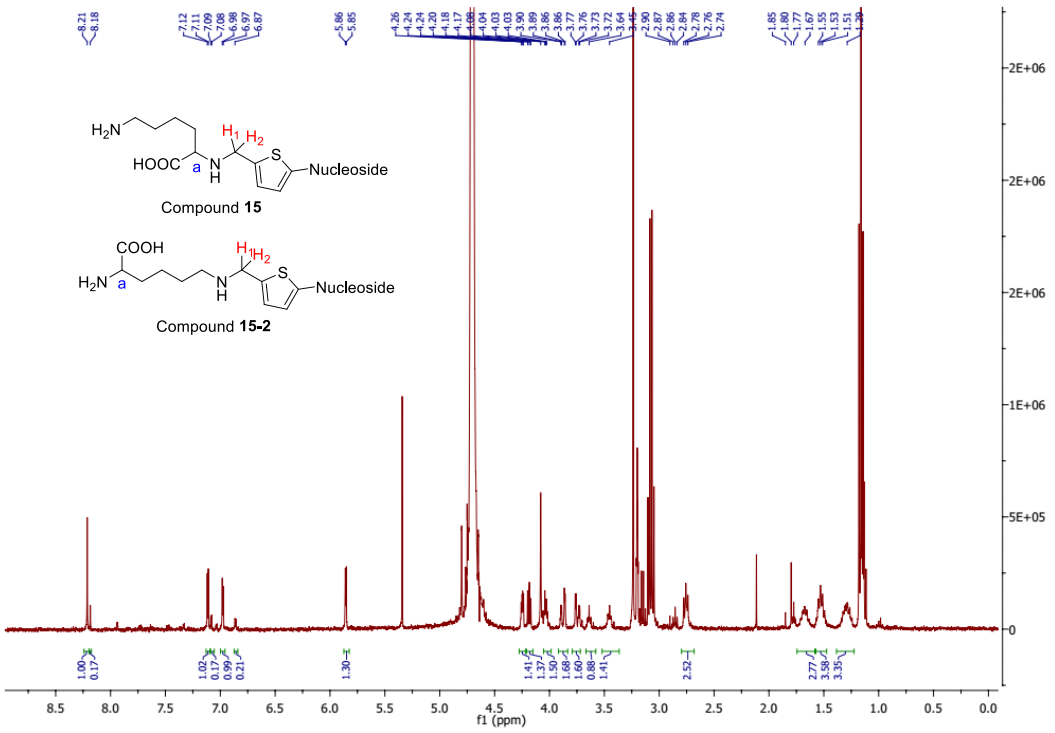


Figure 37 NMR spectra of the mixture of two lysine aminated products **15** and **15-2**

A second reductive aminated nucleoside derivative **16** was prepared in a similar way starting from **8** and Glycine (3 equiv.). After purification of the crude, the 5-FT-Gly uridine **16** was obtained in 27% yield (**entry 2**). Then using the similar condition, the reductive amination of 5-FT uridine with L-valine, L-glutamic acid and L-tryptophan separately were carried out as well, affording the desired uridine analogue **17**, **18** and **19**, respectively (**entry 3**, **4** and **5**). The synthesis of compound **16**, **17**, **18**, **19** were all with a relative low yield, from 20 to 30 %. It is thought that during the C-18 reverse phase chromatography purification, the starting material and amination product could not be separated completely. Therefore, the chromatography purification needed to be carried out more than once, which led to the loss of the products.

Table 4 Reductive amination of 5-FT uridine **8** with different natural amino acids

Entry	Amino acids	Products	Reagents	Yield (%)
1	Lysine	**15**	$NaBH_3CN$, methanol	68
2	Glycine	**16**	$NaBH_3CN$, methanol	27
3	Valine	**17**	$NaBH_3CN$, methanol	22
4	Glutamic acid	**18**	$NaBH_3CN$, methanol	30
5	Tryptophan	**19**	$NaBH_3CN$, methanol	18

2.6. Inhibition assessment of uridine derivatives

In order to assess the inhibition activity of these derivatives, a phosphatase-coupled glycosyltransferase assay was used.[49] This assay uses a calf intestinal alkaline phosphatase (CIP) to cleave inorganic phosphate quantitatively from the nucleoside diphosphate that is formed as a secondary product of the glycosylation reaction. The phosphate concentration is then quantified by a colormetric assay at 620nm. In the presence of inhibitor, the absorbance signal is reduced. The assay can therefore be used to assess the potency of inhibitors by determining their relative IC_{50} values at enzyme turnover rates between 20-50% of UDP-Gal donor. The adapted protocol with this turnover window range could give reproducible and comparable results.

Bovine β-1,4-GalTs, was expressed and purified as previously reported[50] and the protocol used for renaturation of inclusion bodies and refolding of β-1,4-GalTs was adapted from the previous literature.[50] LgtC was expressed and purified as reported.[51] The inhibition assays were carried out as previously reported.[49]

In the first round of screening, uridine derivatives **8**, **10-19** were tested against β-1,4-GalTs under the reported conditions. The maximum concentration of the inhibitor candidates used in the assays was up to 100 μM, while all these derivatives were inactive at all towards β-1,4-GalTs. Then the maximum concentrations of inhibitor candidates were increased to 1 mM. Most of the compounds still showed no activity towards enzyme (**Table 5**), yet compound **13** and **19** exhibited promising activities (**Figure 38**).

Table 5 Inhibitory activities of 5-substituent uridine derivatives **8** and **10** to **19** towards β-1, 4-GalTs

Entry	Cmpd	IC_{50} (μM)	Turnover%
1	**8**	NO inhibition	25 %
2	**10**	NO inhibition	20 %
3	**11**	NO inhibition	28 %
4	**12**	NO inhibition	23 %
5	**13**	207 ± 37	21 %
6	**14**	NO inhibition	29 %
11	**15**	NO inhibition	20 %
12	**16**	NO inhibition	23 %
13	**17**	NO inhibition	26 %
14	**18**	NO inhibition	22 %
15	**19**	290 ± 48	28 %

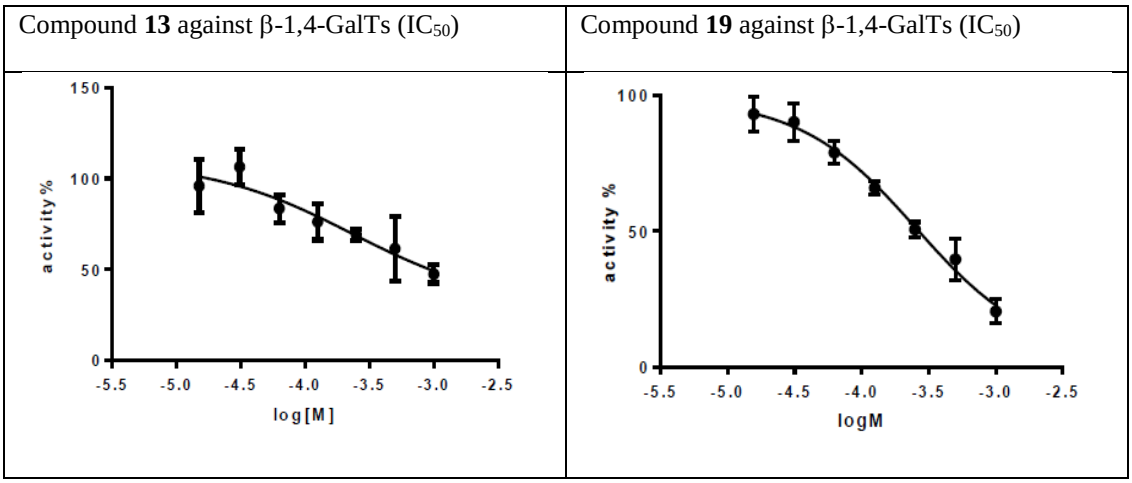
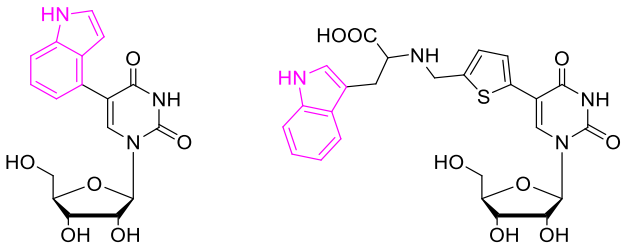


Figure 38 IC_{50} value assessments of compound **13** and **19** against β-1,4-GalTs. Conditions: **13** or **19** (0 to 1mM) β-1,4-GalTs (diluted to the required concentrations), acceptor (GlcNAc, 5 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (CEL, 1 mg/mL), calf-intestinal alkaline phosphate (CIP, 10 U/mL), DMSO (10%) buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 °C with shaking for 20 min. The reaction was stopped by the addition of malachite green reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments

Analysis of the structures of the two active analogues shows that the 5-substituents in both cases contain an indole motif (**Figure 39**). This suggested that the indole substituent in position 5 might provide good affinity towards enzyme active site.



5-indole uridine **13**
IC_{50} = 207 ± 37 μM

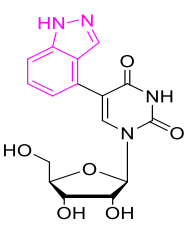
5-FT-Typ uridine **19**
IC_{50} = 290 ± 48 μM

Figure 39 Active nucleoside derivatives **13** and **19** which both contain indole moiety in 5-substituent, with IC_{50} value between 200 to 300 μM.

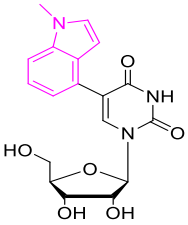
In order to further investigate this possibility, another four derivatives containing an indole or indazole moiety were designed and synthesized (**Table 6**). It was speculated that the

reactivity of these boronic acids might be similar as the indole-4-boronic acid, therefore the Suzuki-cross coupling reaction was tested under the previous experimental conditions. The reaction of 5-iodo uridine and indazole-4-boronic acid in the present of $PdNa_2Cl_4$ (5 mol%), Cs_2CO_3 (2 equiv.) and TPPTs (0.0625 equiv.) at MW 120 ^oC over 30min resulted to the target compound **20** in a 10 % yield (**Table 6**, **entry1**). The yield obtained using this reaction condition was lower than the one obtained for compound **13**. It was thought that the reactivity was dependant on the nature of the boronic acid. In the case of compound **21**, the cross-coupling reaction was conduct between 1-methyl-indole-4-boronic acid and 5-iodo uridine in the same condition. In contrast to the result of synthesis compound **13**, little product was observed with a 4% yield (**entry 2**). In order to improve the yield, other condition was attempted. The mixture of 5-iodo uridine and boronic acid (1.5 equiv.) in the presence of ($PdCl_2(dppf)CH_2Cl_2$ (5 mol%) and $NaHCO_3$ (3 equiv.) in DME/H_2O (3;1) at MW 130 ^oC for 30 min afforded the product **22** in 58 % yield (**entry 3**). Then the cross-couplings were performed between 5-iodo uridine and indole-5-boronic acid and indole-6-boronic acid respectively, giving the corresponding product **23** and **24** in 70% and 64 % yield separately.

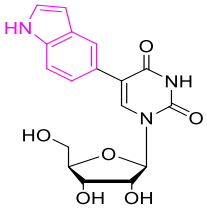
Table 6 Suzuki-Miyaura coupling for synthesis of uridine derivatives with indole or indazole moieties



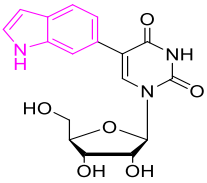
2g



2h



2i



2j

Entry	R	Cmpd	Reagents	MW	Yield (%)
1	5-(1H-indazole)-4-yl	**20**	$PdCl_2$(dppf)DCM, $NaHCO_3$, DME/H_2O (3:1)	MW 120 ºC, 0.5h	10
2	5-(1-methyl-indole)-4-yl	**21**	$PdNa_2Cl_4$, TPPTs, Cs_2CO_3, H_2O	MW 120 ºC, 0.5h	4
3	5-(1-methyl-indole)-4-yl	**21**	$PdCl_2$(dppf)DCM, $NaHCO_3$, DME/H_2O (3:1)	MW 130 ºC, 0.5h	58
4	5-indole-5-yl	**22**	$PdNa_2Cl_4$, TPPTs, Cs_2CO_3, H_2O	MW 120 ºC, 0.5h	70
5	5-indole-6-yl	**23**	$PdNa_2Cl_4$, TPPTs, Cs_2CO_3, H_2O	MW 120 ºC, 0.5h	64

With sufficient quantities of these derivatives in hand, the phosphatase-coupled glycosyltransferase assay was carried out to assess their inhibitory activities and the results were illustrated in **Table 7**.

Table 7 Inhibitory activity of 5-substituent uridine derivatives **20** to **23** towards β-1, 4-GalTs

Entry	Cmpd	IC_{50} (μM)	Turnover%
1	**20**	NO inhibition	42%
2	**21**	163 ± 30	32%
3	**22**	NO inhibition	47%
4	**23**	250 ± 24	34%

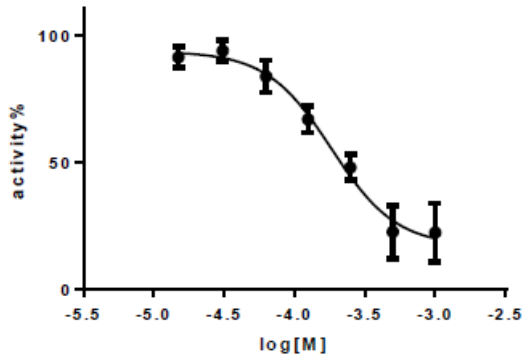
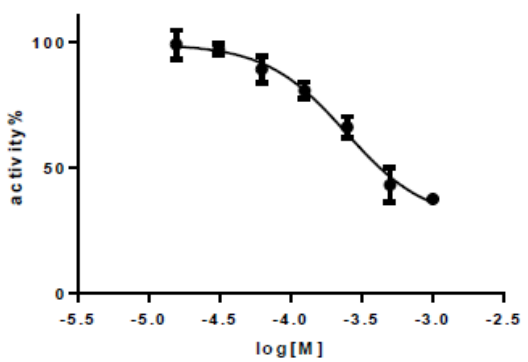
Compound **21** against β-1,4-GalTs (IC_{50})	Compound **23** against β-1,4-GalTs (IC_{50})
	

Figure 40 Inhibition assay of indole containing uridine derivatives **21** and **23** towards β-1,4-GalTs. Conditions: **2.19** or **2.21**(0 to 1mM) β-1,4-GalTs (diluted to the required concentrations), acceptor (GlcNAc, 5 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (CEL, 1 mg/mL), calf-intestinal alkaline phosphate (CIP, 10 U/mL), DMSO (10%) buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 ºC with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments

Encouragingly, compounds **21** and **23** show comparable or a little better activity than compounds **13** and **19**, with IC_{50} values of 163 μM and 250 μM respectively (**Figure 40**). This result could support the hypothesis that the indole substituent in position 5 of uridine is indeed advantageous for inhibitory activity in this series.

The used colorimetric assay detects the formation of the secondary product of β-1,4-GalTs reaction, uridine diphosphate (UDP). The active nucleoside derivatives may also interfere with the phosphatase or chemically react with other assay reagents to confound the assay readout by producing apparent inhibitory activity, leading to false positive results. Therefore, the control experiments were carried out to eliminate this possibility. The control assay excluding β-1,4-GalTs was carried out in a separate plate to identify this possible interference.

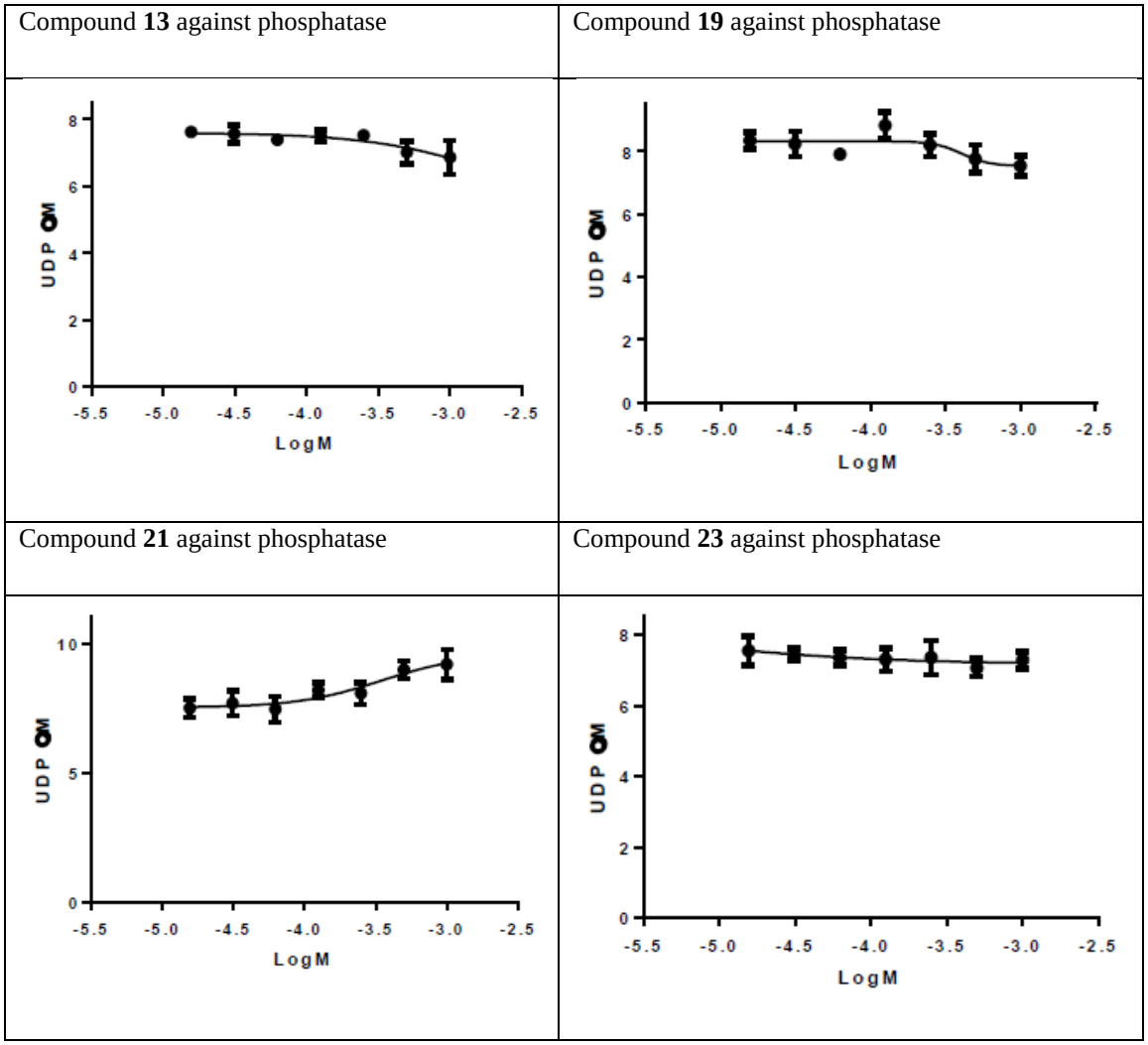


Figure 41 Control experiment of active nucleoside derivatives towards phosphatase Conditions: compound **13**, **19**, **21** or **23** (0 to 1mM), acceptor (GlcNAc, 5 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (CEL, 1 mg/mL), calf-intestinal alkaline phosphate (CIP, 10 U/mL), DMSO (10%), buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 ºC with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments

The active derivatives, **13**, **19**, **21** and **23**, did not interfere with the colorimetric readout of the assay, as evidenced by the lack of signal drop in the control experiment, carried out with increasing concentration of derivatives but without β-1,4-GalTs (**Figure 41**). Based on the previous obtained results, it was confirmed that nucleoside derivatives **13**, **19**, **21** and **23** could show inhibitory activity towards β-1,4-GalTs.

2.7. Synthesis of sugar nucleotide derivatives.

From the previous experimental results, it was suggested that the indole substituent in 5 position of nucleoside was indeed advantageous for inhibitory activity towards enzyme. Given the excellent inhibitory activities that uridine derivatives with 5-substituent containing indole moiety exhibited, there is a good reason to expect the high affinities that corresponding full UDP-nucleotide derivatives would show because of the additional binding activities from 5-position substituents (**Figure 42**). And also, in order to better understand the respective contribution of the 5-substituent, pyrophosphate bond and sugar groups towards enzyme inhibition, the UDP-Gal derivatives corresponding to nucleoside derivatives, 5-indole UDP-Gal and 5-formylthien-tryptophan UDP-Gal (5-FT-Typ UDP-Gal), were prepared.

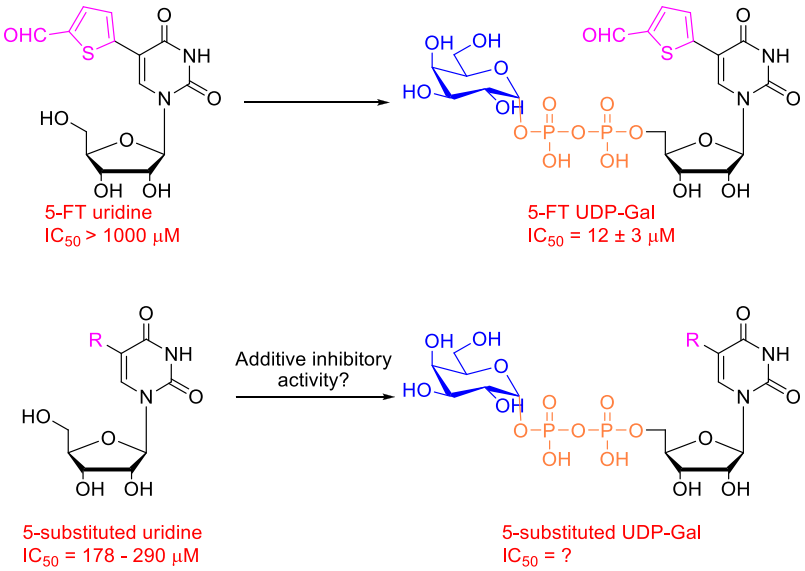
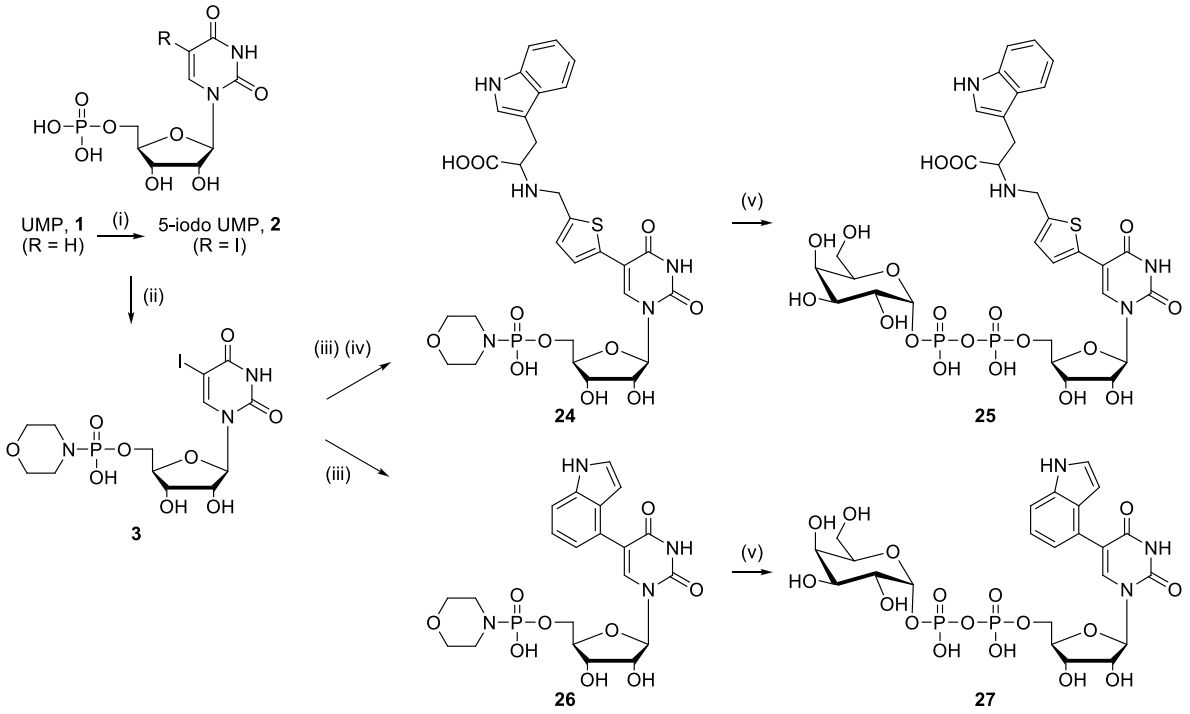


Figure 42 Design of 5-substituted sugar nucleotides corresponding to the most active uridine-based inhibitors

According to the previous synthetic experiments, harsh reaction conditions were required for introduction of substituents in 5-position, yet the pyrophosphate bond is unstable. In addition,

it required 24 hours for reductive reaction to accomplish, which was also impossible for pyrophosphate bond. Therefore, it was planned to introduce 5-position substituent prior to the coupling with sugar moiety, which was different from the previous synthetic route developed by our group. (**Scheme 7**)



Scheme 7 Synthesis of UDP-sugar derivatives. Reagents & conditions: (i) I_2, 2M HNO_3, $CHCl_3$, 18h, 90 °C, 75 %; (ii) Morpholine, dipyridyldisulfide, PPh_3, DMSO, rt, 1.5h 78 %; (iii) Indole-4-boronic acid (for indole UDP-Gal) or 5-formylthien-2-yl-boronic acid (for 5-FT-Typ UDP-Gal), Cs_2CO_3, TPPTs, Na_2PdCl_4, degassed H_2O, MW 120 °C, 30 min, 44 % **(26)** or 47 % (intermediate of 5-FT-Typ UMP morpholidate); (iv) L-Tryptophan, $NaBH_3CN$, degassed methanol, rt, overnight, 40 % (**24**); (v) α-D-galactose-1-phosphate, 2-methylisothiazolone hydrochloride, dry DMF, rt, 9h / 38 % (**27**) or 15h / 35 % (**25**).

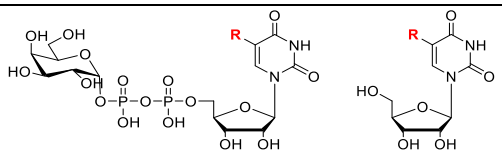
Firstly, the iodination of uridine mono-phosphate was carried out under relatively harsh conditions (80 °C, 2M HNO_3) and 5-iodo UMP **2** hate was obtained after 12 h in 75 % yield. The subsequent phosphormorpholidation was performed under Mukaiyama conditions. After stirring at rt for 2h, the sodium iodide solution in acetone was added to precipitate the desired product. Both indole and 5-formylthien-2-yl groups were introduced at the 5-position of uracil respectively *via* microwave Suzuki cross-coupling and the reductive amination was

carried out with 5- formylthien-2-yl uridine phosphormorpholidate. With the catalysis of NMICl, the phosphormorpholidate product **26** was coupled with α-D-galactose-1-phosphate, reacting at rt for 9h with a yield of 38%. As for the reaction between **24** and α-D-galactose-1-phosphate, longer reaction time, 15h, was used. It was speculated that the steric hindrance from the bulky substitutent in 5-position may slow down coupling reaction with α-D-galactose-1-phosphate.

2.8. Assessment of inhibitory activities of UDP-nucleotides

After obtaining the desired UDP-sugar derivatives, the phosphatase-coupled glycosyltransferase assay was utilized for assessing the potency of their inhibition and the results were illustrated in **Table 8**. As anticipated, compared to corresponding uridine derivatives, the full UDP-sugars were 5-fold more active and the IC_{50} values towards β-1,4-GalTs were improved obviously, from around 200-300 μM to 40-60 μM. However, the rank order within the UDP-Gal series was not as expected. The inhibitory activities of compounds **25** and **27** were even slightly weaker than the previous prototype inhibitor 5-FT UDP-Gal, even though the corresponding uridine derivatives were more active than 5-FT uridine.

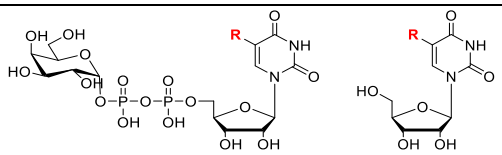
Table 8 Inhibitory activities of 5-substituent uridine derivatives and corresponding 5-substitue UDP-nucleotides towards β-1,4-GalTs.



R	IC_{50} (μM)	IC_{50} (μM)
Indole-4-yl	41 ± 12	207 ± 37
FT-Typ-	60 ± 3	290 ± 48
FT-	12 ± 3	NO inhibition

In order to assess their target selectivity, the active nucleosides, **13** and **19**, and corresponding UDP-nucleotides derivatives, **27** and **25**, were evaluated against other GalTs, LgtC from *N. meningitides*. The results were shown in **Table 9**. All the active uridine derivatives towards β-1,4-GalTs were all inactive against LgtC. In contrast to the nucleosides, the corresponding UDP-nucleotides exhibited similar inhibitory activities against LgtC.

Table 9 Inhibitory activity of 5-substituent uridine derivatives and corresponding5-substitue UDP-nucleotides towards LgtC.



R	IC_{50} (μM)	IC_{50} (μM)
Indole-4-yl	54 ± 13	NO inhibition
FT-Typ-	42 ± 8	NO inhibition
FT-	Not tested	NO inhibition

.

The donor analogues, compound **25** and **27,** acted as inhibitors towards both two GalTs. In order to assess their substrate activity towards both β-1,4-GalTs and LgtC, the phosphatase-coupled glycosyltransferase assays were carried out with these derivatives respectively instead of the natural donor UDP-Gal as donor. (**Figure 43**)

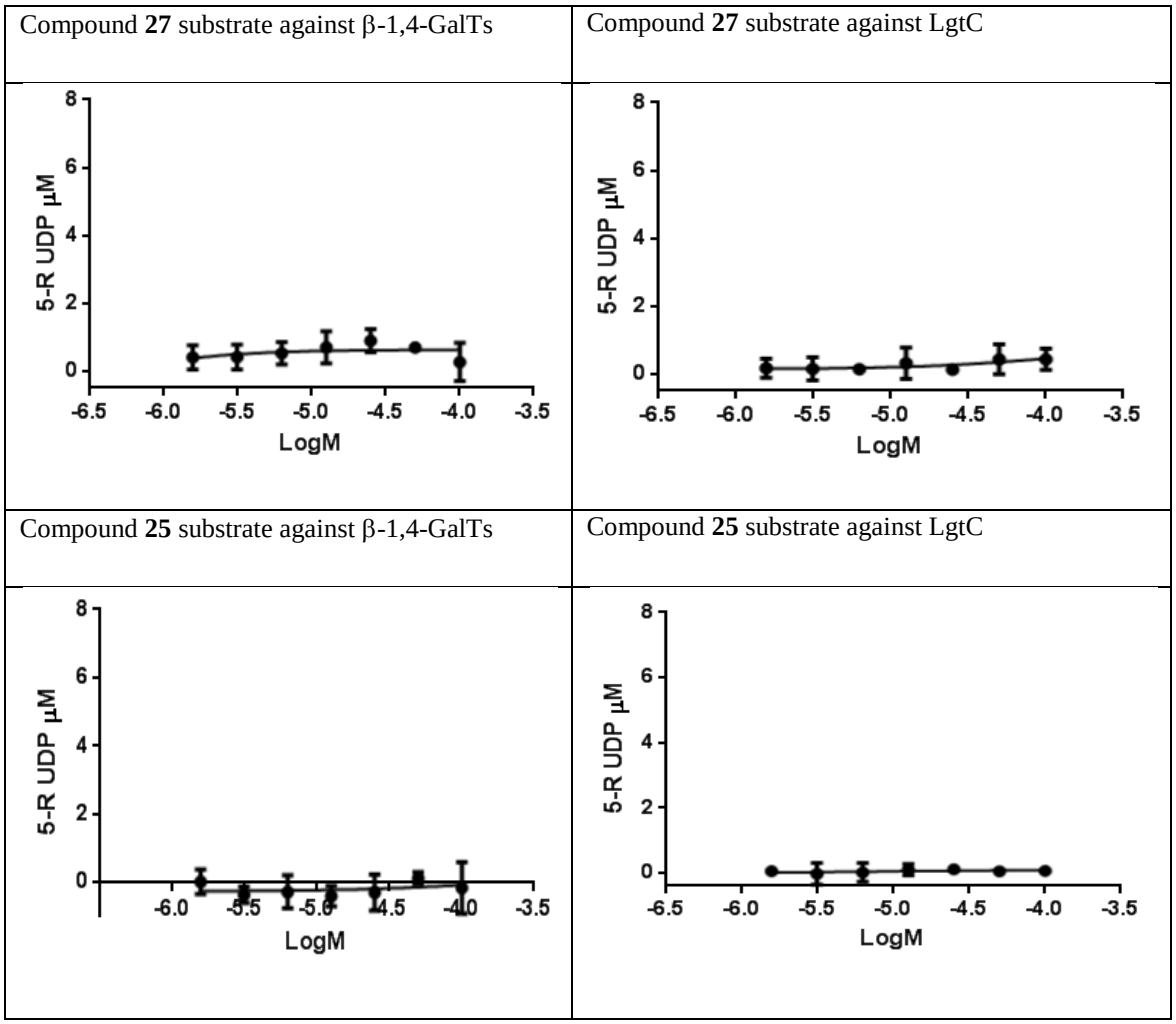


Figure 43 Substrate assay of compound **25** and **27** towards β-1,4-GalTs and LgtC. Conditions: **25** or **27** (0 to 1mM), enzyme (β-1,4-GalTs or LgtC, diluted into the required concentration), acceptor (GlcNAc, 5 mM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (CEL, 1 mg/mL), calf-intestinal alkaline phosphate (CIP, 10 U/mL), DMSO (10%) buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 ^{o}C with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments

Based on these results, it was speculated that the contribution towards β-1,4-GalTs inhibition of the substituent in 5-position and the pyrophosphate bond and α-D-galactose groups was

not additive. In contrast, it appeared to be a direct interplay for GalTs inhibition between different moieties. In the presence of the complete UDP-nucleotide structure, the pyrophosphate group appears to be the dominant factor for enzyme inhibition and enables strong inhibition, even if the nature of the 5-substituent is sub-optimal (**Figure 44**). On the other side, the binding between pyrophosphate bond and manganese ion is quite strong that leads to the potent binding affinity between enzyme and derivatives. Despite significant structural differences between these GalTs, all of them have flexible active site loops folding once binding with the sugar-nucleotides. The 5-substituent could interfere with the movement of the loop once derivatives binding with enzyme, showing inhibition activities. Thus, the selectivity of these derivatives was limited, as evidenced by the indiscriminate activity of UDP-nucleotides against both β-1,4-GalTs and LgtC.

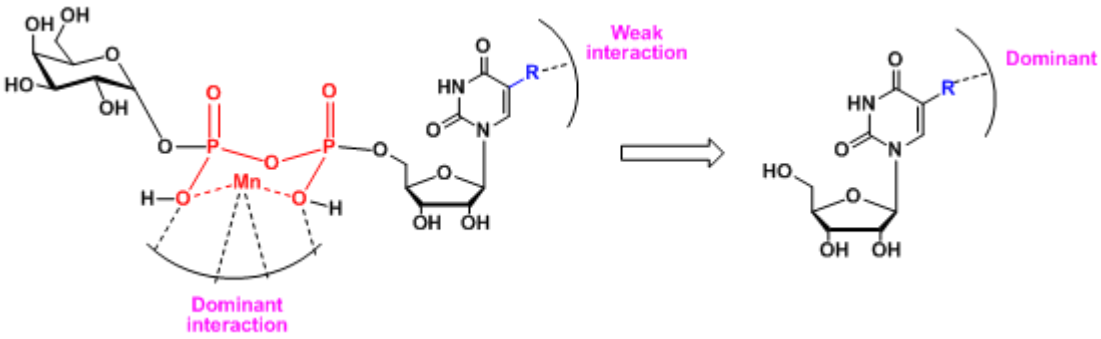


Figure 44 Hypothetical model for the relative contribution of the 5-substituent and the pyrophosphate and α-D-galactose moieties towards enzyme inhibition

In contrast, in the absence of the pyrophosphate bond and galactose groups, the nature of the 5-position substituent becomes essential not only for this mode of inhibition in general, but also for inhibitory potency. In the case of β-1,4-GalTs, indole-containing 5-position substituents appear to be favourable, probably due to some specific interaction with the target enzyme. For other GalTs, like LgtC, the nature of these interactions may well be different,

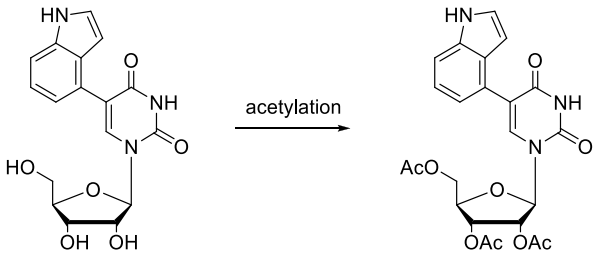
due to the different architecture of the flexible loop and C-terminal region in different enzymes. This model therefore suggests that 5-substitued uridine inhibitors provide an opportunity for the design of target selective GalT inhibitors.

2.9. Cell assay of uridine derivatives

As described above, β-1,4-GalTs catalyse the transfer of galactose from UDP-Gal donor to a terminal GlcNAc/Glucose residue on the acceptor. β-1,4-GalTs are involved in the biosynthesis of PSGL-1, not only of the saccharide epitope sLe^x on PSGL-1, but also the LacNAc linker, the linkage between sLe^x and protein backbone.[49] PSGL-1 acts as a key mediator in the process of inflammatory cell recruitment. As 5-indole uridine **13** exhibited inhibitory activity towards β-1,4-GalTs in the enzyme assay, it was then utilized in the cell assay. The effects of the β-1,4-GalTs inhibitor 5-indole uridine **13** and its acetylated **28** on the cell surface levels of PSGL-1 on human peripheral blood mononuclear cells (hPBMCs) were investigated. These experiments were carried out by Dr Varsha Kanabar in the Sackler Institute of Pulmonary Pharmacology at King's College London. PSGL-1 levels were studies both under basal conditions and upon stimulation of hPBMCs with interleukin-1β *via* an established FACS-based assay.[52] The average cell surface levels of PSGL-1 per cell, which was indicated by mean fluorescence intensity, MFI, was used to quantified the variation of PSGL-1 levels on cell surface.

The concentration of **13** was used from 1 to 100 μM in the assay and the results are shown in **Figure 45**. Even up to the maximum concentration 100 μM, compound **13** still did not affect the cell surface levels on both the 1L-1β-stimulated and basal levels of cell surface PSGL-1.

It was speculated that the membrane permeability of **13** was still poor even in the absence of the pyrophosphate linkage and sugar moiety. As the nucleoside still contained several free hydroxyl groups, this relative polar compound was difficult to penetrate cell membrane. The acetylation reaction of **13** was carried out for the synthesis of pro-drug **28** (**Scheme 8**). Thus, this fully acetylated product was utilized in the cell assay directly as the protecting groups could be readily removed by intracellular carboxylesterases once taken up into the cells.



Scheme 8 Synthesis of acetylated 5-indole uridine. Reagents and conditions: acetic anhydride, 4-Dimethylaminopyridine (DMAP), pyridine, rt, 4h, 77%.

At the concentration of 100 μM, **28** could selectively decrease the cell surface PSGL-1 expression on the condition with stimulation of IL-1β, but showed inactive towards PSGL-1 expression level on the basal condition. (**Figure 45**) These results demonstrated that this compound was taken up into the cells and converted into the parent compound which was catalysed by intracellular carboxylesterases. The inhibitory activity may come from this active nucleoside. In the cell assay profile, it was hypothesized that inhibitors reduce the cell surface PSGL-1 levels not by altering the structure of sLe^x on PSGL-1 but by affecting glycan-dependant trafficking of PGSL-1 to the cell surface. Thus, inhibitors can discriminate between 1L-1β-induced and basal levels of cell surface PSGL-1.

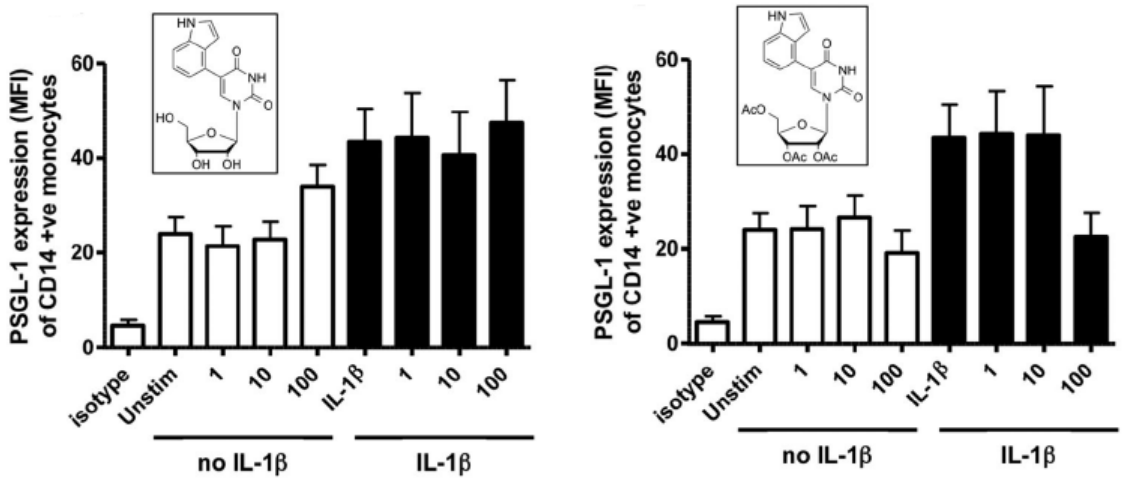


Figure 45 Effects of **13** and its acetylated product **28** on cell surface levels of PSGL-1 on both basal condition and on the stimulation of IL-1β.

2.10. Summary and conclusion

In conclusion, nucleoside-based inhibitors towards β-1,4-GalTs were developed. Amongst these analogues, compounds containing indole moiety in 5-position substituents showed potent activities. The comparison of the corresponding complete UDP-sugar derivatives allowed the identification of structural factors that contributed to inhibitory activity. The effects of the most active nucleoside derivative **13** and its ester prodrug **28** in a PSGL-1 expression assay were also investigated. Compound **28** could selectively decrease the cell surface PSGL-1 expression on the IL-1β-stimulated cells.

2.11. Experimental section.

General. All chemical reagents were obtained commercially and used as received, unless stated otherwise. Microwave-assisted reactions were carried out on a Monowave 300 microwave synthesis reactor from Anton Paar. Thin layer chromatography (TLC) was performed on pre-coated plates of Silica Gel 60 F254 (Merck), with *i*-PA:H_2O:NH_3 = 6:3:1 as the mobile phase, unless otherwise stated. Spots were visualised under UV light (254/365nm). NMR spectra were recorded at 300 K on a Bruker BioSpin machine at, respectively, 400.13 MHz (^{1}H-NMR), 100.62 MHz (^{13}C-NMR) and 161 MHz (^{31}P-NMR). Prior to the recording of ^{31}P-NMR spectra, a drop of triethylamine was added to each sample to suppress line broadening and enhance resolution. Chemicals shifts (δ) are reported in ppm (parts per million) and coupling constants (J) in Hz. Mass spectra were recorded at the EPSRC National Mass Spectrometry Facility in Swansea. All yields (%) are isolated yields.

Column chromatography. Preparative reverse-phase chromatography was performed on a Biologic LP chromatography system equipped with a peristaltic pump and a 254 nm UV Optics Module under the following conditions: Ion-pair chromatography was performed using Lichroprep RP-18 resin equilibrated with 0.05 M TEAB (triethylammonium bicarbonate, pH 7.3). Gradient: 0 – 30% MeOH against 0.05 M TEAB over a total volume of 400 mL. Flow rate: 2 mL/min. Product-containing fractions were combined and repeatedly co-evaporated with methanol to remove residual TEAB.

Suzuki cross-coupling reaction

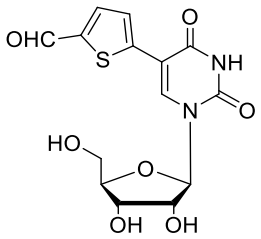
General method A: 5-iodouridine (1 equiv.), boronic acid (1.5 equiv.) and Cs_2CO_3 (2 equiv.) in degassed water were combined in a sealable microwave tube under nitrogen atmosphere.

Na_2PdCl_4 (0.025 equiv.) and TPPTs (0.0625 equiv.) were added. The vessel was sealed and the mixture was heated in the microwave reactor at 120°C for 30min. After cooling to room temperature, the solvent was evaporated and the reaction purified by column chromatography.

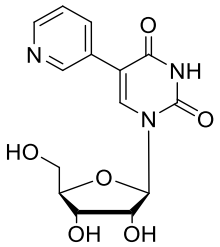
General method B. 5-iodouridine (1 equiv.) and boronic acid (1.5 equiv.) in degassed DME/water (3:1) were combined in a sealable microwave tube under nitrogen atmosphere. $PdCl_2$(dppf)DCM (0.05 equiv.) and $NaHCO_3$ (3 equiv.) were added. The vessel was sealed and the mixture was heated in the microwave reactor at 130°C for 30min. After cooling to room temperature, the solvent was evaporated and the reaction purified by column chromatography.

General method C. 5-iodouridine (1 equiv.) and boronic acid (1.5 equiv.) in degassed dioxane/water (3:1) were combined in a sealable microwave tube under nitrogen atmosphere. $PdCl_2$(dppf)DCM (0.05 equiv.) and $NaHCO_3$ (3 equiv.) were added. The vessel was sealed and the mixture was heated in the microwave reactor at 130°C for 30min. After cooling to room temperature, the solvent was evaporated and the reaction purified by column chromatography.

General method D: 5-iodouridine (1 equiv.) and boronic acid (1.5 equiv.) were combined in degassed DME in a sealable microwave tube under nitrogen atmosphere. $PdCl_2$(dppf)DCM (0.05 equiv.) and K_2CO_3 (2 equiv.) were added. The vessel was sealed and the mixture was heated in the microwave reactor at 130°C for 30min. After cooling to room temperature, the solvent was evaporated and the reaction purified by column chromatography.

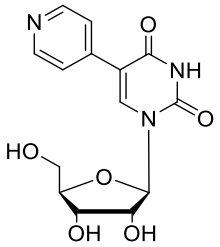


5-(1,2,3,4-tetrahydro-1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl) furan-2-yl)-2,4-dioxopyrimidin-5-yl)thiophene-2-carbaldehyde (8). Compound **8** was synthesised from 5-iodo uridine (50 mg, 0.135 mmol) and 5-formyl-2-thiophene-boronic acid *via* general method A, and obtained as a white powder in a yield of 54% (25.5 mg). ^{1}H-NMR (400 MHz, MeOD) 3.85 (1H, dd, *J* = 2, 12 Hz), 4.03 (1H, dd, *J* = 2.4, 12 Hz), 4.11 (1H, m), 4.27 (2H, m), 6.00 (1H, d, *J* = 3.2 Hz), 7.63 (1H, d, *J* = 4 Hz), 7.84 (1H, d, *J* = 4 Hz), 9.12 (1H, s), 9.86 (1H, s). ^{13}C-NMR (100 MHz, d_6-DMSO) 59.5 (C-5'), 68.6 (C-3'), 74.4 (C-2'), 84.3 (C-4'), 89.4 (C-1'), 106.7 (C-5), 122.7, 137.1, 138.5, 141.4, 149.2 (C-6 + C-Thiophene), 144.2 (C-2), 161.2 (C-4), 184.1 (CHO). m/z (ESI) 355.0627 $[M+H]^+$, $C_{14}H_{15}N_2O_7S$ requires 355.0600.

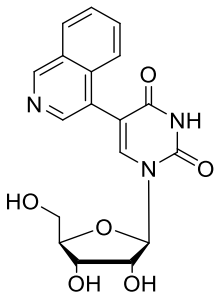


1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(pyridin-3-yl) pyrimidine-2,4(1H,3H)-dione (10). Compound **10** was synthesised from 5-iodo uridine (20 mg, 0.054 mmol) and 3-pyridineboronic acid *via* general method B, and obtained as a white powder in a yield of 20% (4.1 mg). ^{1}H-NMR (400 MHz, MeOD) 3.80 (1H, dd, *J* = 2.4, 12 Hz), 3.92 (1H, dd, *J* = 2.8, 12 Hz), 4.07 (1H, m), 4.27 (2H, m), 5.99 (1H, d, *J* = 3.2 Hz),

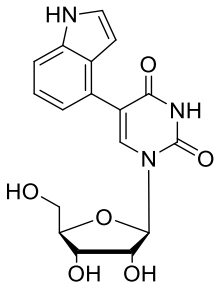
7.47 (1H, dd, *J* = 5.2, 8 Hz), 8.08 (1H, dt, *J* = 2, 8 Hz), 8.48 (1H, dd, *J* = 1.6, 5.2 Hz), 8.59 (1H, s), 8.78 (1H, d, *J* = 2.4 Hz). ^{13}C-NMR (100 MHz, MeOD) 62.0 (C-5'), 70.8 (C-3'), 76.3 (C-2'), 86.2 (C-4'), 91.2 (C-1'), 112.2 (C-5), 124.9, 131.4, 137.8, 140.8, 140.9, 148.7 (C-6 + C-Pyridine), 149.3 (C-2), 152.0 (C-4). m/z (ESI) 322.2961 $[M+H]^+$, $C_{14}H_{16}N_3O_6$ requires 322.1039.



1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(pyridin-4-yl)pyrimidine-2,4(1H,3H)-dione (11). Compound **11** was synthesised from 5-iodo uridine (20 mg, 0.054 mmol) and 4-pyridineboronic acid via general method B, and obtained as a light yellow powder in a yield of 27% (5 mg). ^{1}H-NMR (400 MHz, D_2O) 3.66 (1H, dd, *J* = 5.6, 12.8 Hz), 3.78 (1H, dd, *J* = 2.8, 12.8 Hz), 3.99 (1H, m), 4.28 (1H, t, *J* = 6 Hz), 4.38 (1H, m), 5.75 (1H, d, *J* = 3.2 Hz), 6.49 (2H, d, *J* = 7.6 Hz), 7.80 (1H, d, *J* = 6.8 Hz), 7.94 (1H, s). ^{13}C-NMR (100 MHz, d_6-DMSO) 60.1 (C-5'), 69.3 (C-3'), 73.9 (C-2'), 84.7 (C-4'), 91.3 (C-1'), 112.5 (C-5), 128.0, 128.6 (*m*-Py, *m*-Py), 145.1 (C-6), 146.0 (*p*-Py), 148.1 (C-2), 150.1, 150.3 (*o*-Py, *o*-Py), 160.5 (C-4). m/z (ESI) 322.2951 $[M+H]^+$, $C_{14}H_{16}N_3O_6$ requires 322.1039.

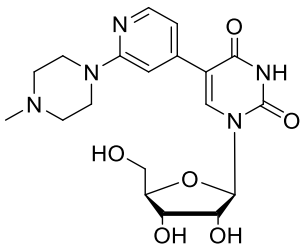


1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(isoquinolin-4-yl)pyrimidine-2,4(1H,3H)-dione (12). Compound **12** was synthesised from 5-iodo uridine (20 mg, 0.054 mmol) and 4-isoquinolineboronic acid *via* general method A, and obtained as a light yellow powder in a yield of 25% (4.7 mg). ^{1}H-NMR (400 MHz, MeOD), 3.67 (1H, dd, *J* = 2.8, 12Hz), 3.79 (1H, dd, *J* = 2.8, 12 Hz), 4.05 (1H, m), 4.22 (1H, t, *J* = 5.2 Hz), 4.32 (1H, t, *J* = 4.8 Hz), 6.04 (1H, d, *J* = 4.4 Hz), 7.73 (1H, dt, *J* = 1.2, 6.8 Hz), 7.81 (1H, dt, *J* = 1.2, 7.6 Hz), 7.87 (1H, m), 8.17 (1H, dd, *J* = 1.2, 8.4 Hz), 8.37 (1H, s), 8.39 (1H, s), 9.26 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 61.8 (C-5'), 71.1 (C-3'), 76.1 (C-2'), 86.4 (C-4'), 91.0 (C-1'), 111.9 (C-5), 126.1, 127.0, 127.1, 129.1, 129.3, 129.9, 132.4, 136.8, 142.5, 143.8 (C-6 + isoquinoline), 152.7 (C-2), 153.8 (C-4). m/z (ESI) 372.1290 $[M+H]^{+}$, $C_{18}H_{18}N_3O_6$ requires 372.1196.

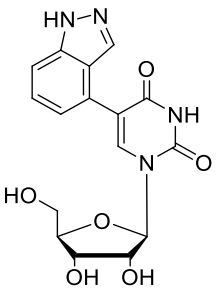


1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(1H-indol-4-yl) pyrimidine-2,4(1H,3H)-dione (13). Compound **13** was synthesised from 5-iodo uridine (60 mg, 0.162 mmol) and indole-4-boronic acid *via* general method A, and obtained as a white powder in a yield of 43% (18 mg). ^{1}H-NMR (400 MHz, MeOD) 3.67 (1H, dd, *J* = 3.6, 12 Hz), 3.77 (1H, dd, *J* = 4, 12 Hz), 4.04 (1H, m), 4.18 (1H, t, *J* = 5.6Hz), 4.32 (1H, t, *J* = 5.2 Hz), 6.06 (1H, d, *J* = 5.2 Hz), 6.42 (1H, d, *J* = 3.2 Hz), 7.08 (1H, d, *J* = 1.2, 7.2 Hz), 7.17 (1H, t, *J* = 7.6 Hz), 7.27 (1H, d , *J* = 3.2 Hz), 7.40 (1H, d, *J* = 8 Hz), 8.23 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 62.4 (C-5'), 71.2 (C-3'), 75.8 (C-2'), 86.5 (C-4'), 90.6 (C-1'), 101.8 (C-indole-

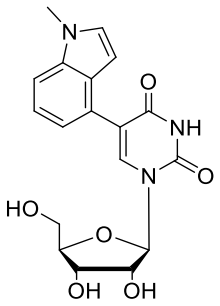
3), 112.2 (C-5), 116.5, 121.4, 122.1, 125.7, 125.9, 128.5, 137.9, 140.6 (C-6 + C-indole), 152.5 (C-2), 164.9 (C-4). m/z (ESI) 360.1193 $[M+H]^+$, $C_{17}H_{18}N_3O_6$ requires 360.1196.



1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(2-(4-methyl piperazin-1-yl)pyridin-4-yl)pyrimidine-2,4(1H,3H)-dione (14). Compound **14** was synthesised from 5-iodo uridine (20 mg, 0.054 mmol) and 2-(4-methylpiperzin-1-yl)pyridine-4-boronic acid *via* general method B, and obtained as a white powder in a yield of 35% (8 mg). ^{1}H-NMR (400 MHz, MeOD) 2.36 (3H, s), 2.60 (4H, m), 3.57 (4H, m), 3.80 (1H, dd, *J* = 2.4, 12 Hz), 3.93 (1H, dd, *J* = 2.4, 12 Hz), 4.08 (1H, m), 4.7 (1H, m), 5.99 (1H, d, *J* = 3.6 Hz), 6.97 (1H, dd, *J* = 1.2, 5.6Hz), 7.19 (1H, s), 8.07 (1H, d, *J* = 1.2, 5.6Hz), 8.64 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 45.6 (C-NMe), 46.1, 46.3 (C-piperazin-3, C-piperazin-5), 55.7, 55.9 (C-piperazin-2, C-piperazin-6), 61.6 (C-5'), 70.9 (C-3'), 76.4 (C-2'), 86.3 (C-4'), 91.2 (C-1'), 107.9 (C-pyridine), 113.3 (C-5), 113.8, 141.5, 144.4, 148.3, 151.8 (C-6 + C-pyridine), 161.2 (C-2), 164.0 (C-4). m/z (ESI) 420.1911 $[M+H]^+$, $C_{19}H_{26}N_5O_6$ requires 420.1883.

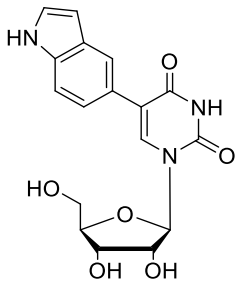


1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(1H-indazol-4-yl) pyrimidine-2,4(1H,3H)-dione (20). Compound **20** was synthesized from 5-iodo uridine (60 mg, 0.152 mmol) and 1H-indazole-4-boronic acid *via* general method D, and obtained as a light grey powder in a yield of 10% (6 mg). ^{1}H-NMR (400 MHz, MeOD) 3.70 (1H, dd, *J* = 2.4, 12 Hz), 3.82 (1H, dd, *J* = 2.8, 12 Hz), 4.06 (1H, m), 4.22 (1H, t, *J* = 4.8 Hz), 4.33 (1H, t, *J* = 4.4 Hz), 6.06 (1H, d, *J* = 4.8 Hz), 7.22 (1H, d, *J* = 8 Hz), 7.43 (1H, dd, *J* = 8.8, 7.2 Hz), 7.54 (1H, d, *J* = 8.4 Hz), 8.01 (1H, s), 8.44 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 62.1 (C-5'), 71.5 (C-3'), 76.1 (C-2'), 86.5 (C-4'), 90.7 (C-1'), 110.9 (C-5), 115.0 (C-indazole-3), 122.6, 123.3, 127.5, 127.8, 127.9, 135.0, 141.2 (C-6 + C-indazole), 152.3 (C-2), 164.5 (C-4). m/z (ESI) 361.1146 $[M+H]^{+}$, $C_{16}H_{17}N_4O_6$ requires 361.1148.

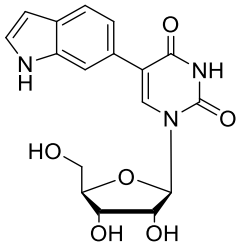


1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(1-methyl-1H-indol-4-yl)pyrimidine-2,4(1H,3H)-dione (21). Compound **21** was synthesized from 5-iodo uridine (60 mg, 0.152 mmol) and 1-methyl-1H-indole-4-boronic acid pinacol ester *via* general method C, and obtained as a white powder in a yield of 58% (35 mg). ^{1}H-NMR (400 MHz, MeOD) 3.58 (1H, m), 3.69 (3H, s), 3.70 (1H, dd, *J* = 3.2, 12.4 Hz), 3.98 (1H, m), 4.08 (1H, t, *J* = 5.6 Hz), 4.23 (1H, t, *J* = 4 Hz), 5.85 (1H, d, *J* = 4 Hz), 6.23 (1H, d, *J* = 4 Hz), 6.96 (1H, d, *J* = 7.2 Hz), 7.16 (1H, d, *J* = 3.2 Hz), 7.18 (1H, d, *J* = 7.2 Hz), 7.38 (1H, d, *J* = 8.4 Hz), 7.92 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 47.8 (C-NMe), 62.4 (C-5'), 71.8 (C-3'), 75.9 (C-2'), 86.5 (C-4'), 90.5 (C-1'), 101.1 (C-indole-3), 110.9 (C-5), 116.1, 121.5, 122.2, 126.0,

129.0, 130.3, 138.5, 140.8 (C-6 + C-indole), 152.4 (C-2), 164.9 (C-4). m/z (ESI) 374.1350 $[M+H]^+$, $C_{18}H_{20}N_3O_6$ requires 374.1352.



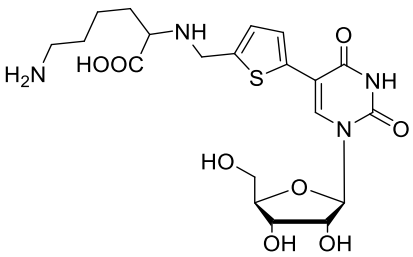
1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(1H-indol-5-yl)pyrimidine-2,4(1H,3H)-dione (22). Compound **22** was synthesized from 5-iodo uridine (60 mg, 0.152 mmol) and indole-5-boronic acid *via* general method A, and obtained as a light grey powder in a yield of 70% (45 mg). ^{1}H-NMR (400 MHz, MeOD) 3.75 (1H, dd, *J* = 2.8, 12 Hz), 3.86 (1H, dd, *J* = 2.8, 12 Hz), 4.06 (1H, m), 4.24 (1H, t, *J* = 4.8 Hz), 4.32 (1H, t, *J* = 4.8 Hz), 6.04 (1H, d, *J* = 4.8 Hz), 6.47 (1H, s), 7.25 (1H, d, *J* = 2.8 Hz), 7.27 (1H, dd, *J* = 1.2, 4.4 Hz), 7.40 (1H, d, *J* = 8.4 Hz), 7.75 (1H, s), 8.22 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 62.2 (C-5'), 71.4 (C-3'), 75.9 (C-2'), 86.2 (C-4'), 91.1 (C-1'), 102.7 (C-indole-3), 111.9 (C-5), 118.3, 121.5, 123.4, 125.7, 126.1, 129.5, 137.2, 138.8 (C-6 + C-indole), 154.7 (C-2), 168.5 (C-4). m/z (ESI) 360.1194 $[M+H]^+$, $C_{17}H_{18}N_3O_6$ requires 360.1196.



1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-5-(1H-indol-6-yl)pyrimidine-2,4(1H,3H)-dione (23). Compound **23** was synthesized from 5-iodo uridine (60 mg, 0.152 mmol) and indole-6-boronic acid *via* general method A, and obtained as a white powder in a yield of 64% (40 mg). ^{1}H-NMR (400 MHz, MeOD) 3.76 (1H, dd, *J* = 2.8, 12 Hz), 3.87 (1H, dd, *J* = 2.8, 12 Hz), 4.07 (1H, m), 4.24 (1H, t, *J* = 4.8 Hz), 4.31 (1H, t, *J* = 4.8 Hz), 6.03 (1H, d, *J* = 4.4 Hz), 6.44 (1H, d, *J* = 4 Hz), 7.17 (1H, dd, *J* = 5.6, 8.4 Hz), 7.25 (1H, d, *J* = 3.2 Hz), 7.55 (1H, dd, *J* = 8.4 Hz), 7.66 (1H, s), 7.28 (1H, s). ^{13}C-NMR (100MHz, MeOD) 62.1 (C-5'), 71.4 (C-3'), 76.0 (C-2'), 86.3 (C-4'), 90.7 (C-1'), 102.3 (C-indole-3), 112.5 (C-5), 117.5, 120.7, 121.0, 126.5, 126.9, 129.1, 137.6, 139.3 (C-6 + C-indole), 152.3 (C-2), 165.2 (C-4). m/z (ESI) 360.1194 $[M+H]^+$, $C_{17}H_{18}N_3O_6$ requires 360.1196.

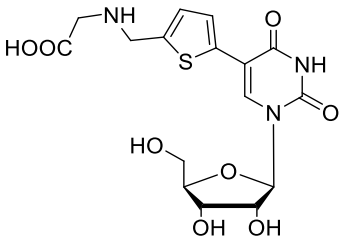
Reductive amination reactions: compounds 15-19

General protocol. Under nitrogen, the respective amino acid (1 equiv.) was added to a stirred solution of **8** (1 equiv.) in degassed methanol. After stirring for 1h, $NaBH_3CN$ (0.1 equiv.) in methanol was added through a syringe. After stirring for 24h, the solvent was removed under reduced pressure and the residue was purified by reverse phase column chromatography (0.05M TEAB against methanol, gradient: 0%-20%).

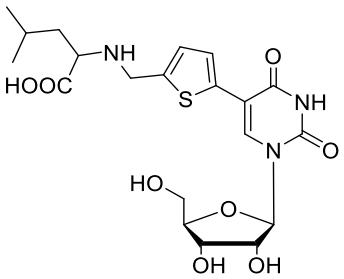


2-((5-(1,2,3,4-tetrahydro-1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl)furan-2-yl)-2,4-dioxopyrimidin-5-yl)thiophen-2-yl)methylamino)-6-aminohexanoic acid

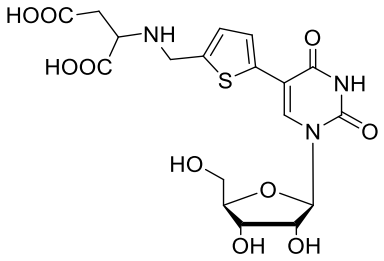
(2.13). Compound **15** was synthesized from **8** (10 mg, 0.028 mmol) and L-Lysine, and obtained as a light yellow powder in a yield of 68% (9.2 mg). ^{1}H-NMR (400 MHz, D_2O) 1.30 (2H, m), 1.54 (2H, m), 1.68 (2H, m), 2.77 (2H, m), 3.75 (1H, dd, *J* = 3.2, 12.8 Hz), 3.9 (1H, dd, *J* = 2.4, 12.8 Hz), 4.03 (1H, m), 4.08 (1H, m), 4.18 (1H, t, *J* = 5.6 Hz), 4.25 (1H, dd, *J* = 3.2, 5.2 Hz), 5.34 (1H, s), 5.85 (1H, d, *J* = 3.2 Hz), 6.98 (1H, d, *J* = 3.6 Hz), 8.21 (1H, s). ^{13}C-NMR (100 MHz, D_2O) 22.1 (C-Lysine-4), 27.2 (C-Lysine-5), 30.2 (C-Lysine-3), 42.2 (C-Lysine-6), 46.2 (C-NH-CH_2-), 60.0 (C-Lysine-2), 67.7 (C-5'), 68.7 (C-3'), 74.2 (C-2'), 80.7 (C-4'), 81.5 (C-1'), 112.4 (C-5), 118.6, 124.0, 130.4, 130.7, 131.9 (C-6 + C-thiophene), 150.5 (C-2), 160.8 (C-4), 174.6 (C-COOH). m/z (ESI) 483.1697 $[M-H]^-$, $C_{20}H_{27}N_4O_8S$ requires 483.1550.



2-((5-(1,2,3,4-tetrahydro-1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl) furan-2-yl)-2,4-dioxopyrimidin-5-yl)thiophen-2-yl)methylamino) acetic acid (16). Compound **16** was synthesised from **8** (11 mg, 0.031 mmol) and L-Glycine, and obtained as a light yellow powder in a yield of 27% (3.5 mg). ^{1}H-NMR (400 MHz, D_2O) 3.26 (2H, s), 3.78 (1H, dd, *J* = 4.4, 12.8 Hz), 3.90 (1H, dd, *J* = 4.4, 12 Hz), 4.07 (1H, m), 4.21 (1H, m), 4.27 (1H, m), 4.40 (2H, s), 5.85 (1H, d, *J* = 2.8 Hz), 7.14 (1H, d, *J* = 4 Hz), 7.20 (1H, d, *J* = 3.6 Hz), 8.40 (1H, s). ^{13}C-NMR (100 MHz, D_2O) 51.7 (C-NHCH_2-), 57.8 (C-glycine-2), 61.7 (C-5'), 71.0 (C-3'), 76.4 (C-2'), 86.3 (C-4'), 91.1 (C-1'), 110.1 (C-5), 124.2, 127.6, 136.0, 137.4, 141.7 (C-6 + C-thiophene), 151.7 (C-2), 161.5 (C-4), 170.4 (C-COOH). m/z (ESI) 412.1789 $[M-H]^-$, $C_{16}H_{18}N_3O_8S$ requires 412.0815.

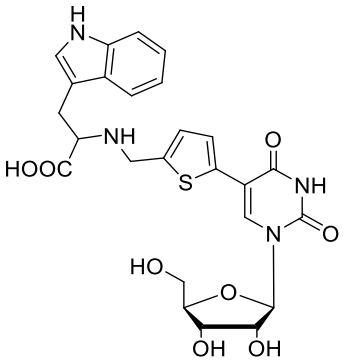


2-((5-(1,2,3,4-tetrahydro-1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl) furan-2-yl)-2,4-dioxopyrimidin-5-yl)thiophen-2-yl)methylamino)-4-methylpentanoic acid (17). Compound **17** was synthesised from **8** (10 mg, 0.028 mmol) and L-Leucine, and obtained as a white powder in 22% yield (3 mg). ^{1}H-NMR (400 MHz, MeOD) 0.90 (3H, d, *J* = 6.4 Hz), 0.96 (3H, d, *J* = 6.4 Hz), 1.45 (1H, m), 1.56 (1H, m), 1.80 (1H, m), 3.82 (1H, dd, *J* = 2, 8 Hz), 3.90 (1H, m), 3.96 (1H, dd, *J* = 2, 8 Hz), 4.28 (2H, m), 6.00 (1H, d, *J* = 3.6 Hz), 6.97 (1H, d, *J* = 3.6 Hz), 7.34 (1H, d, *J* = 3.6 Hz), 8.59 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 23.1 (C-Me), 23.2 (C-Me), 25.2 (C-CH), 47.6 (C-Leucine-3), 49.5 (C-NH-CH_2), 66.8 (C-Leucine-2), 69.5 (C-5'), 75.6 (C-3'), 77.2 (C-2'), 88.6 (C-4'), 98.7 (C-1'), 116.9 (C-5), 132.8, 135.6, 139.2, 140.0, 144.5 (C-6 + C-thiophene), 155.9 (C-2), 168.2 (C-4), 177.9 (C-COOH). m/z (ESI) 468.1444 $[M-H]^-$, $C_{20}H_{26}N_3O_8S$ requires 468.1441.



2-((5-(1,2,3,4-tetrahydro-1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl) furan-2-yl)-2,4-dioxopyrimidin-5-yl)thiophen-2-yl)methylamino)-3-(1H-indol-3-yl) propanoic acid (18). Compound **18** was synthesised from **8** (10 mg, 0.028 mmol) and L-Glutamic acid, and obtained as a light yellow powder in 18% yield (2.5 mg). ^{1}H-NMR (400

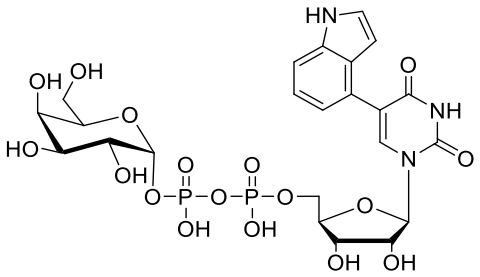
MHz, MeOD) 0.92 (2H, m), 1.23 (2H, m), 2.50 (1H, dd, *J* = 2, 12 Hz), 2.62 (1H, dd, *J* = 2, 12 Hz), 2.67 (1H, m), 2.75 (1H, m), 2.93 (2H, m), 3.28 (2H, m), 4.66 (1H, d, *J* = 2.8 Hz), 5.63 (1H, d, *J* = 3.6 Hz), 6.00 (1H, d, *J* = 3.6 Hz), 7.38 (1H, s). ^{13}C-NMR (100 MHz, MeOD) 38.7 (C-Glu-3), 48.6 (C-NHCH$_2$), 61.5 (C-5'), 64.0 (C-Glu-2), 70.3 (C-3'), 73.5 (C-2'), 84.7 (C-4'), 98.6 (C-1'), 111.2 (C-5), 122.7, 126.0, 134.4, 135.9, 140.2 (C-6 + C-thienphene), 150.0 (C-2), 161.2 (C-4), 174.2 (C-COOH), 178.5 (C-COOH). m/z (ESI) 484.0966 [M-H]$^-$, $C_{19}H_{22}N_3O_{10}S$ requires 484.1026.



2-((5-(1,2,3,4-tetrahydro-1-((2R,3S,4R,5R)-tetrahydro-3,4-dihydroxy-5-(hydroxymethyl) furan-2-yl)-2,4-dioxopyrimidin-5-yl)thiophen-2-yl)methylamino)-3-(1H-indol-3-yl)propanoic acid (19). Compound **19** was synthesised from **8** (8 mg, 0.023 mol) and L-Tryptophan, and obtained as a light yellow powder in 30% yield (3.7 mg). ^{1}H-NMR (400 MHz, D$_2$O) 2.92 (1H, m), 3.42 (1H, t, *J* = 6.8 Hz), 3.74 (2H, m), 3.88 (2H, m), 4.06 (1H, dd, *J* = 2.8, 6.4 Hz), 4.19 (1H, dd, *J* = 5.2, 6 Hz), 4.27 (1H, dd, *J* = 5.2, 3.6 Hz), 5.87 (1H, d, *J* = 3.6 Hz), 6.71 (1H, d, *J* = 3.6 Hz), 6.88 (1H, t, *J* = 8 Hz), 6.93 (1H, d, *J* = 8.4 Hz), 6.98 (1H, t, *J* = 8.4 Hz), 7.07 (1H, s), 7.28 (1H, d, *J* = 8 Hz), 7.39 (1H, d, *J* = 8 Hz), 8.08 (1H, s). ^{13}C-NMR (100 MHz, D$_2$O) 32.3 (C-Trp-3), 47.3 (C-NHCH$_2$), 64.7 (C-Trp-2), 63.2 (C-5'), 76.6 (C-3'), 77.1 (C-2'), 85.9 (C-4'), 91.3 (C-1'), 108.7 (C-indole-3), 111.9 (C-5), 113.1, 115.8, 119.0, 122.1, 122.9, 126.9, 127.5, 129.8, 131.4, 137.5, 139.2, 148.5 (C-6 + C-thiophene + C-

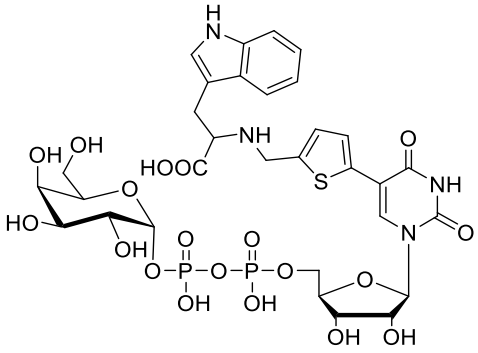
indole), 155.9 (C-2), 169.1 (C-4), 177.9 (C-COOH). m/z (ESI) 541.1393 $[M-H]^-$, $C_{25}H_{26}N_4O_8S$ requires 541.1393.

Synthesis of sugar-nucleotides. 5-iodo UMP **2** and its phosphoromorpholidate **3** were synthesized as previously reported. The 5-substituent derivatives **24** and **26** were obtained by Suzuki cross-coupling of 5-iodo UMP phosphoromorpholidate **3** with, respectively, 5-formyl-2-thiopheneboronic acid or indole-4-boronic acid, followed in the case of **24** by reductive amination. Phosphoromorpholidates **24** and **26** were freeze-dried overnight and added separately to added with α-D-Galactose-1-phosphate and NMICl in the dry DMF. The reaction was stirred at room temperature for 9 h (**27**) or 15h (**25**). All solvents were removed under reduced pressure, and the crude product was purified by anion-exchange chromatography and reverse phase column. Product-containing fractions were combined and reduced to dryness. After repeated co-evaporation of the residue with methanol, extra TEAB was removed.

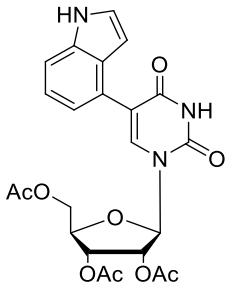


Compound **27** was obtained in a yield of 38%, white powder. ^{1}H-NMR (400 MHz, D_2O) 1.12 (9H, t, *J* = 7.8 Hz), 3.04 (6H, q, *J* = 7.2 Hz), 3.46 (1H, dd, *J* = 4.8, 12 Hz), 3.54 (1H, dd, *J* = 7.6, 12 Hz), 3.56 (1H, m), 3.68 (1H, m), 3.77 (1H, d, *J* = 4.8 Hz), 3.92 (1H, dd, *J* = 4.8, 7.6 Hz), 4.00 (2H, m), 4.14 (1H, m), 4.20 (1H, m), 4.31 (1H, m), 5.40 (1H, m), 5.91 (1H, d, *J* = 6 Hz), 6.27 (1H, d, *J* = 2.8 Hz), 7.01 (1H, d, *J* =7.2 Hz), 7.15 (1H, t, *J* = 7.6 Hz), 7.28 (1H, d, *J* = 2.8 Hz), 7.43 (1H, d, *J* = 8.4 Hz), 7.73 (1H, s). ^{13}C-NMR (100 MHz, D_2O) 60.9 (C-6”),

65.0 (d, $J_{c,p}$ = 10 Hz, C-5'), 69.1 (C-2''), 69.8 (C-3''), 70.0 (C-4''), 71.8 (C-4'), 73.0 (C-5''), 75.0 (C-3'), 76.2 (C-2'), 76.8 (C-5''), 88.6 (C-4'), 95.6 (C-1'), 95.7 (d, $J_{c,p}$ = 10 Hz, C-1''), 100.1 (C-indole-3), 112.1 (C-5), 115.8, 120.6, 121.8, 123.5, 126.1, 126.4, 135.7, 139.1 (C-6 + C-indole), 151.6 (C-2), 164.5 (C-4). ^{31}P-NMR (161 MHz, D_2O) -11.2 (d, *J* = 21.2Hz), -13.6 (d, *J* = 21.2 Hz). m/z (ESI) 680.0885 $[M-H]^-$, $C_{23}H_{29}N_3O_{17}P_2$ requires 681.0972.



Compound **2.25** was obtained in a yield of 35%, light yellow powder. ^{1}H-NMR (400 MHz, D_2O) 1.10 (9H, t, *J* = 7.2 Hz), 3.00 (6H, q, *J* = 7.2 Hz), 3.48 (2H, m), 3.55 (1H, dd, *J* = 3.2, 8 Hz), 3.6 (1H, m), 3.72 (3H, m), 3.94 (1H, m), 3.96 (1H, m), 4.13 (2H, m), 4.17 (1H, m), 4.26 (1H, t, *J* = 4 Hz), 4.32 (1H, t, *J* = 4 Hz), 5.49 (1H, dd, *J* = 4, 8 Hz), 5.93 (1H, d, *J* = 5.2 Hz), 6.72 (1H, d, *J* = 4 Hz), 6.84 (1H, t, *J* = 8 Hz), 6.94 (1H, t, *J* = 8 Hz), 7.10 (1H, s), 7.25 (1H, d, *J* = 7.6 Hz), 7.37 (1H, d, *J* = 8 Hz), 7.76 (1H, s). ^{13}C-NMR (100 MHz, D_2O) 27.2 (C-Typ-3), 60.4 (C-6''), 65.5 (d, $J_{c,p}$ = 10 Hz, C-5'), 66.9 (C-Typ-2), 68.3 (C-3''), 69.1 (C-2''), 69.3 (C-4''), 69.9 (C-5''), 71.9 (C-3'), 73.7 (C-2'), 83.7 (C-4'), 95.8 (C-1'), 100.9 (d, $J_{c,p}$ = 10 Hz, C-1''), 107.7 (C-indole-3), 111.8 (C-5), 118.3, 119.2, 121.6, 124.4, 124.9, 126.1, 130.0, 130.0, 134.6, 135.8, 135.9, 136.5 (C-6 + C-thienphene + C-indole), 150.6 (C-2), 163.8 (C-4), 170.5 (C-COOH). ^{31}P-NMR (161 MHz, D_2O) -11.3 (d, *J* = 19.3Hz), -12.8 (d, *J* = 19.3 Hz). m/z (ESI) 865.0885 $[M-H]^-$, $C_{23}H_{29}N_3O_{17}P_2$ requires 864.0972.



A solution of 5-indole uridine **8** (54 mg, 0.15 mmol) and DMAP (0.9 mg, 0.075 mmol) was dissolved in 0.5 mL pyridine and acetic anhydride (54 mg, 0.52 mmol) was added. The reaction was stirred for 4 h at rt before quenched with water. The solvent was removed by evaporation and co-evaporated three times with toluene. The light yellow crude compound was purified by flash chromatography (Hexane / EA = 2: 1 to 1 :1) to yield acetylated compound **28** as grey solid (58 mg, 77 %). ^{1}HNMR (400 Hz, $CDCl_3$) 1.79 (3H, s), 2.04 (3H, s), 2.07 (3H, s), 4.30 (3H, m), 5.30 (1H, dd, *J* = 4, 6 Hz), 5.39 (1H, t, *J* = 6 Hz), 6.14 (1H, d, *J* = 6 Hz), 6.43 (1 H, m), 7.03 (1H, dt, *J* = 1.2, 8 Hz), 7.09 (1H, m), 7.52 (1H, m), 7.58 (1H, s), 8.51 (1H, s). ^{13}C-NMR (100 MHz, $CDCl_3$) 20.1 (C-Me), 20.5 (C-Me), 20.6 (C-Me), 63.3 (C-5'), 70.4 (C-3'), 72.7 (C-2'), 80.1 (C-4'), 86.9 (C-1'), 102.3 (C-indole-3), 111.6 (C-5), 117.3, 119.6, 120.7, 125.2, 125.5, 127.9, 135.8, 135.9 (C-6 + C-indole), 150.0 (C-2), 162.3 (C-4), 169.8 (C-C=O), 170.4 (C-C=O), 171.0 (C-C=O). m/z (ESI) 486.1320 $[M+H]^-$, $C_{23}H_{23}N_3O_9$ requires 485.1434.

GalT inhibition experiments.

General. All reagents for the biochemical assays were obtained commercially and used as received, unless otherwise stated. Bovine β-1, 4-GalTs was expressed, purified and refolded using an adaptation of a previously reported protocol.[50] LgtC was expressed and purified as

previously reported.[51] Inhibition assays were carried out as previously reported.[49] Absorbance measurements were carried out on a BMG Labtech POLARstar Optima multiplate reader.

Data collection and analysis protocol. Assays were carried out on 96-well plates. On each microplate, sample, control and background wells were included in triplicate. A calibration curve (0-12.5 μM UDP, corresponding to 0-25 μM Pi) was constructed for each microplate by linear regression. The calibration curve was used to convert absorbance measurements at 620 nm in sample and control wells to [UDP] (μM). For each sample and control well, a corresponding background well (containing identical components but no acceptor) was included, to account for non-specific hydrolysis of donor. Corrected absorbance values for each well were obtained by subtracting the corresponding background reading form the absorbance of the respective sample or control well. Inhibition (%) was calculated by dividing absorbance in the presence of inhibitor by maximum absorbance (negative control, no inhibitor). Percentage inhibition was plotted over log[inhibitor] and analysed with GraphPad Prism 6 software to obtain relative IC_{50} values. Averages and standard deviations were calculated in Microsoft Excel.

PSGL-1 expression assay This cell assay was carried out as reported.[52] Peripheral venous blood was collected from healthy donors (n =5). Mononuclear cells were separated by density after centrifugation step. The monocyte layers were then washed with media, which was followed by adding cells into each well of a 96-well plate. The compound **8** and ester prodrug **28** at 1, 10, or 100 μM were added for 1h, followed by either media or 10 ng/ mL 1L-1β in the continued presence of compounds for 72 h at 37 ºC and 5 % CO_2. Cells were harvested and analysed by flow cytometry.

2.12. References

1. Zawilska, J. B.; Wojcieszak, J.; Olejniczak, A. B., Prodrugs: A challenge for the drug development. *Pharmacol Rep* **2013,** *65*, 1-14.
2. Stella, V. J., Prodrugs: Some Thoughts and Current Issues. *J Pharm Sci-Us* **2010,** *99*, 4755-4765.
3. Schultz, C., Prodrugs of biologically active phosphate esters. *Bioorgan Med Chem* **2003,** *11*, 885-898.
4. Wiemer, A. J.; Wiemer, D. F., Prodrugs of Phosphonates and Phosphates: Crossing the Membrane Barrier. *Top Curr Chem* **2015,** *360*, 115-160.
5. Krise, J. P.; Stella, V. J., Prodrugs of phosphates, phosphonates, and phosphinates. *Adv Drug Deliver Rev* **1996,** *19*, 287-310.
6. Metcalf, W. W.; van der Donk, W. A., Biosynthesis of Phosphonic and Phosphinic Acid Natural Products. *Annu Rev Biochem* **2009,** *78*, 65-94.
7. Wiemer, D. F., Synthesis of nonracemic phosphonates. *Tetrahedron* **1997,** *53*, 16609-16644.
8. Huttunen, K. M.; Rautio, J., Prodrugs - An Efficient Way to Breach Delivery and Targeting Barriers. *Curr Top Med Chem* **2011,** *11*, 2265-2287.
9. Ghosh, S.; Chan, J. M. W.; Lea, C. R.; Meints, G. A.; Lewis, J. C.; Tovian, Z. S.; Flessner, R. M.; Loftus, T. C.; Bruchhaus, I.; Kendrick, H.; Croft, S. L.; Kemp, R. G.; Kobayashi, S.; Nozaki, T.; Oldfield, E., Effects of bisphosphonates on the growth of Entamoeba histolytica and Plasmodium species in vitro and in vivo. *J Med Chem* **2004,** *47*, 175-187.
10. Hecker, S. J.; Erion, M. D., Prodrugs of phosphates and phosphonates. *J Med Chem* **2008,** *51*, 2328-45.
11. Maria, E. A., Current Prodrug Strategies for the Delivery of Nucleotides into Cells. *Drug Design Rev - Online (Discontinued)* **2005,** *2*, 373-387.
12. Declercq, E., Biochemical Aspects of the Selective Antiherpes Activity of Nucleoside Analogs. *Biochem Pharmacol* **1984,** *33*, 2159-2169.
13. Kogler, M.; Vanderhoydonck, B.; De Jonghe, S.; Rozenski, J.; Van Belle, K.; Herman, J.; Louat, T.; Parchina, A.; Sibley, C.; Lescrinier, E.; Herdewijn, P., Synthesis and Evaluation of 5-Substituted 2 '-deoxyuridine Monophosphate Analogues As Inhibitors of Flavin-Dependent Thymidylate Synthase in Mycobacterium tuberculosis. *J Med Chem* **2011,** *54*,

4847-4862.

14. Declercq, E.; Descamps, J.; Verhelst, G.; Walker, R. T.; Jones, A. S.; Torrence, P. F.; Shugar, D., Comparative Efficacy of Antiherpes Drugs against Different Strains of Herpes-Simplex Virus. *J Infect Dis* **1980,** *141*, 563-574.

15. De Clercq, E., The Unabated Synthesis of New Nucleoside Analogues with Antiviral Potential: A Tribute to Morris J. Robins. *Nucleos Nucleot Nucl* **2009,** *28*, 586-600.

16. Jordheim, L. P.; Durantel, D.; Zoulim, F.; Dumontet, C., Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat Rev Drug Disc* **2013,** *12*, 447-464.

17. Turek, P.; Kotora, M.; Hocek, M.; Votruba, I., Novel method for preparation of highly substituted 6-arylpurines by reactions of 6-alkynylpurines with zirconacyclopentadienes. *Collect Czech Chem C* **2005,** *70*, 339-349.

18. Szombati, Z.; Baerns, S.; Marx, A.; Meier, C., Synthesis of C8-Arylamine-Modified 2'-Deoxyadenosine Phosphoramidites and their Site-Specific Incorporation into Oligonucleotides. *Chembiochem* **2012,** *13*, 700-712.

19. Saito, Y.; Miyamoto, S.; Suzuki, A.; Matsumoto, K.; Ishihara, T.; Saito, I., Fluorescent nucleosides with 'on-off' switching function, pH-responsive fluorescent uridine derivatives. *Bioorg Med Chem Lett* **2012,** *22*, 2753-2756.

20. Baba, S.; Negishi, E., Novel Stereospecific Alkenyl-Alkenyl Cross-Coupling by a Palladium-Catalyzed or Nickel-Catalyzed Reaction of Alkenylalanes with Alkenyl Halides. *J Am Chem Soc* **1976,** *98*, 6729-6731.

21. Heck, R. F.; Nolley, J. P., Palladium-Catalyzed Vinylic Hydrogen Substitution Reactions with Aryl, Benzyl, and Styryl Halides. *J Org Chem* **1972,** *37*, 2320-&.

22. Dieck, H. A.; Heck, F. R., Palladium Catalyzed Synthesis of Aryl, Heterocyclic and Vinylic Acetylene Derivatives. *J Organomet Chem* **1975,** *93*, 259-263.

23. Miyaura, N.; Yamada, K.; Suzuki, A., New Stereospecific Cross-Coupling by the Palladium-Catalyzed Reaction of 1-Alkenylboranes with 1-Alkenyl or 1-Alkynyl Halides. *Tetrahedron Lett* **1979,** *20*, 3437-3440.

24. Miyaura, N.; Suzuki, A., Stereoselective Synthesis of Arylated (E)-Alkenes by the Reaction of Alk-1-Enylboranes with Aryl Halides in the Presence of Palladium Catalyst. *J Chem Soc Chem Comm* **1979**, 866-867.

25. Suzuki, A., Carbon-carbon bonding made easy. *Chem Commun* **2005**, 4759-4763.

26. Miyaura, N.; Suzuki, A., Palladium-Catalyzed Cross-Coupling Reactions of Organoboron Compounds. *Chem Rev* **1995,** *95*, 2457-2483.

27. Fuwa, H.; Suzuki, T.; Kubo, H.; Yamori, T.; Sasaki, M., Total synthesis and biological assessment of (-)-exiguolide and analogues. *Chemistry* **2011,** *17*, 2678-88.

28. Kwak, J. H.; Cho, Y. A.; Jang, J. Y.; Seo, S. Y.; Lee, H.; Hong, J. T.; Han, S. B.; Lee, K.; Kwak, Y. S.; Jung, J. K., Expedient synthesis of 4-O-methylhonokiol via Suzuki-Miyaura cross-coupling. *Tetrahedron* **2011,** *67*, 9401-9404.

29. Pellicena, M.; Kramer, K.; Romea, P.; Urpi, F., Total synthesis of (+)-herboxidiene from two chiral lactate-derived ketones. *Org Lett* **2011,** *13*, 5350-3.

30. Theeramunkong, S.; Caldarelli, A.; Massarotti, A.; Aprile, S.; Caprioglio, D.; Zaninetti, R.; Teruggi, A.; Pirali, T.; Grosa, G.; Tron, G. C.; Genazzani, A. A., Regioselective Suzuki Coupling of Dihaloheteroaromatic Compounds as a Rapid Strategy To Synthesize Potent Rigid Combretastatin Analogues. *J Med Chem* **2011,** *54*, 4977-4986.

31. Berezin, A. A.; Koutentis, P. A., Suzuki-Miyaura reactions of the soluble guanylate cyclase inhibitor NS2028: a non-product specific route to C-8 substituted analogues. *Tetrahedron* **2011,** *67*, 4069-4078.

32. Evitt, A.; Tedaldi, L. M.; Wagner, G. K., One-step synthesis of novel glycosyltransferase inhibitors. *Chem Commun* **2012,** *48*, 11856-11858.

33. Rouhi, A. M., A renaissance in organocatalysis. *Chem Eng News* **2004,** *82*, 41-45.

34. Herve, G.; Sartori, G.; Enderlin, G.; Mackenzie, G.; Len, C., Palladium-catalyzed Suzuki reaction in aqueous solvents applied to unprotected nucleosides and nucleotides. *Rsc Adv* **2014,** *4*, 18558-18594.

35. Shaughnessy, K. H., Palladium-Catalyzed Modification of Unprotected Nucleosides, Nucleotides, and Oligonucleotides. *Molecules* **2015,** *20*, 9419-9454.

36. Agrofoglio, L. A.; Gillaizeau, I.; Saito, Y., Palladium-assisted routes to nucleosides. *Chem Rev* **2003,** *103*, 1875-1916.

37. Lakshman, M. K.; Hilmer, J. H.; Martin, J. Q.; Keeler, J. C.; Dinh, Y. Q.; Ngassa, F. N.; Russon, L. M., Palladium catalysis for the synthesis of hydrophobic C-6 and C-2 aryl 2'-deoxynucleosides. Comparison of C-C versus C-N bond formation as well as C-6 versus C-2 reactivity. *J Am Chem Soc* **2001,** *123*, 7779-87.

38. Amann, N.; Wagenknecht, H. A., Preparation of pyrenyl-modified nucleosides via Suzuki-Miyaura cross-coupling reactions. *Synlett* **2002**, 687-691.

39. Collier, A.; Wagner, G. K., Suzuki-Miyaura cross-coupling of unprotected halopurine nucleosides in water-influence of catalyst and cosolvent. *Synthetic Commun* **2006,** *36*, 3713-3721.

40. Storr, T. E.; Strohmeier, J. A.; Baumann, C. G.; Fairlamb, I. J., A sequential direct

arylation/Suzuki-Miyaura cross-coupling transformation of unprotected 2'-deoxyadenosine affords a novel class of fluorescent analogues. *Chem Commun (Camb)* **2010,** *46*, 6470-2.

41. Fresneau, N.; Hiebel, M. A.; Agrofoglio, L. A.; Berteina-Raboin, S., Efficient Synthesis of Unprotected C-5-Aryl/Heteroaryl-2 '-deoxyuridine via a Suzuki-Miyaura Reaction in Aqueous Media. *Molecules* **2012,** *17*, 14409-14417.

42. Pesnot, T.; Jorgensen, R.; Palcic, M. M.; Wagner, G. K., Structural and mechanistic basis for a new mode of glycosyltransferase inhibition. *Nat Chem Biol* **2010,** *6*, 321-323.

43. Descroix, K.; Pesnot, T.; Yoshimura, Y.; Gehrke, S. S.; Wakarchuk, W.; Palcic, M. M.; Wagner, G. K., Inhibition of Galactosyltransferases by a Novel Class of Donor Analogues. *J Med Chem* **2012,** *55*, 2015-2024.

44. Pesnot, T.; Jørgensen, R.; Palcic, M. M.; Wagner, G. K., Structural and mechanistic basis for a new mode of glycosyltransferase inhibition. *Nat Chem Biol* **2010,** *6*, 321-323.

45. Meinhart, A.; Mutschler, H., Use of a new class of nucleotide sugar as antibiotic or cytostatic agent. Google Patents2012.

46. Pesnot, T.; Jorgensen, R.; Palcic, M. M.; Wagner, G. K., Structural and mechanistic basis for a new mode of glycosyltransferase inhibition. *Nat Chem Biol* **2010,** *6*, 321-3.

47. Alhaj Zen, A.; Aylott, J. W.; Chan, W. C., An appraisal of the Suzuki cross-coupling reaction for the synthesis of novel fluorescent coumarin derivatives. *Tetrahedron Lett* **2014,** *55*, 5521-5524.

48. Denmark, S. E.; Ober, M. H., Organosilicon Reagents: Synthesis and Application to Palladium-Catalyzed Cross-Coupling Reactions. *ChemInform* **2004,** *35*, no-no.

49. Tedaldi, L.; Evitt, A.; Goos, N.; Jiang, J.; Wagner, G. K., A practical glycosyltransferase assay for the identification of new inhibitor chemotypes. *MedChemComm* **2014,** *5*, 1193-1201.

50. Boeggeman, E.; Qasba, P. K., Studies on the metal binding sites in the catalytic domain of beta1,4-galactosyltransferase. *Glycobiology* **2002,** *12*, 395-407.

51. Sebaugh, J. L., Guidelines for accurate EC50/IC50 estimation. *Pharma Statis* **2011,** *10*, 128-134.

52. Kanabar, V.; Tedaldi, L.; Jiang, J.; Nie, X.; Panina, I.; Descroix, K.; Man, F.; Pitchford, S. C.; Page, C. P.; Wagner, G. K., Base-modified UDP-sugars reduce cell surface levels of P-selectin glycoprotein 1 (PSGL-1) on IL-1beta-stimulated human monocytes. *Glycobiology* **2016**.

CHAPTER 3

A dynamic combinatorial chemistry approach for the development of β-1,4-galactosyltransferase inhibitors

3.1. Introduction

In this chapter, a brief introduction will be given to the dynamic combinatorial chemistry (DCC) concept. The latter was the description of attempts to apply DCC strategy to the development of potent nucleoside-based inhibitors. This chapter mainly gave an introduction of the dynamic combinatorial library (DCL) generated via active nucleoside derivatives (which was described in chapter 2) and hydrazines/hydrazones. It was utilized for the screening the replacement of pyrophosphate and galactose moieties

3.1.1. General concept of dynamic combinatorial chemistry (DCC)

Dynamic combinatorial chemistry (DCC) is a powerful strategy for the generation of new bioactive molecules formed by reversible reactions of simple building blocks under thermodynamic control. The interconversion of these building blocks involves covalent or non-covalent bonds.[1, 2] The library of these reversibly interconverting building blocks affords dynamic combinatorial libraries (DCLs). All constituents in a DCL are in equilibrium and the distribution of library members is determined by their thermodynamic stability in the DCL.[1] DCC has been implemented in a variety of chemical systems and applied to different targets, including proteins[3], peptides[4], nucleic acids[5], inorganic[6] and organic ions[7].

Protein-directed dynamic combinatorial chemistry offers a strategy for efficient discovery of novel chemical structures for binding to target proteins (**Figure 46**).[3] It requires a library of molecules that can react reversibly with each other under thermodynamic control to generate chemical diversity and these molecules are referred to as ‘building blocks’. The position of

the equilibrium of DCL is influenced by thermodynamic variables of the system. In the case of protein-directed DCC, the distribution of the library building blocks can be caused to re-equilibrium by the addition of a target protein which is referred as 'template'. In the process, the components with the strongest affinity for the target can be amplified and identified by biophysical techniques. These library components can act as a starting point to guide organic synthesis of stable analogues for further development of enzyme inhibitors. Protein-templated DCC was used for the first time in the late 1990's, leading to the identification of inhibitors towards carbonic anhydrase (CA).[8] With the development of analytical techniques, like HPLC, NMR etc., to date, there are plenty of examples for the successful application of protein-templated DCC to develop novel inhibitors for various enzymes, such as Glutathione S-Transferase[9] and Aspartic Protease[10].

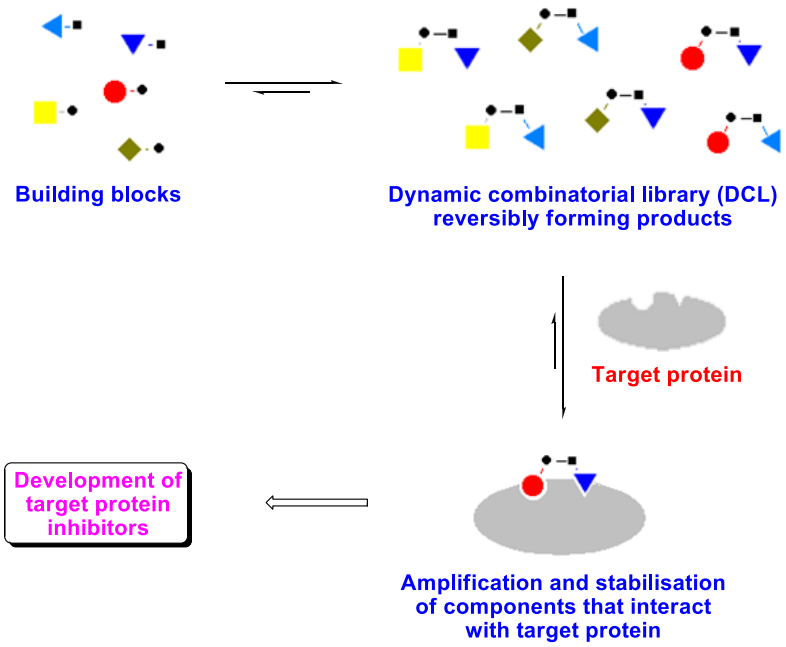


Figure 46 Fundamental concept of protein-templated DCC. Colourful triangles, squares and balls: different structures of building blocks; Black squares and diamonds: functional groups for reversible reactions with other components. (modified from reference[11])

3.1.2. Design of protein-directed DCC experiments

In principle, successful execution of a protein-directed DCC experiment relies on four elements[12]:

(a) A protein template;

(b) A library of suitable building blocks;

(c) A reversible reaction;

(d) A suitable analytical method.

Below, a brief general introduction will be given to these aspects of DCL as this will be essential for the understanding of the design of our experiments.

3.1.2.1. Protein template

The protein in a protein-directed DCC enables to select the most amplified components from the DCL by influencing the equilibrium. It is crucial to ensure that the protein is stable and active under the DCC experimental conditions in a certain pH and temperature window. The harsh conditions, like high temperature and strong acidic or basic condition, may deactive the protein and lead to false negative results. In the case of harsh conditions, use of a catalyst might be required. For example, for a acyl hydrazone DCLs, equilibration between aldehydes and hydrazides requires acidic conditions which are incompatible with target enzyme, aniline was employed as a nucleophilic catalyst which enable DCL formation at biocompatible pH.[13] Also, the amount of enzyme used to re-equilibrate the generation of product needs to be

considered. For instance, a DCL was generated by using imine formation between amines and aldehydes to discover inhibitors of the metalloenzyme human carbonic anhydrase II (hCA II).[14] The experiment was performed in the presence or absence of enzyme (80 μM) and the results were analysed by HPLC. Also, DCC experiment was employed to investigate the active sites of α-chymotrypsin *via* the reversible reaction between boronic acids and diols.[15] The concentration of target enzyme, α-chymotrypsin, was increased to 2 mM (ratio: enzyme/boronic acid = 1/1) as the employed detected technique ^{11}B-NMR is relatively insensitive.

3.1.2.2. Design of dynamic combinatorial library

The design of dynamic combinatorial library is crucial for DCC experiments. In contrast to high throughput screening approaches which use large chemical libraries (typically > 10^6)[16], DCLs require relatively small libraries. The design of DCLs is usually carried out according to the target enzyme structure information. In order to produce DCL efficiently, the building blocks need to fulfil several essential characteristics.[3, 17] Firstly, they must contain functional groups for reversible reactions with other components in the library. Secondly, they are usually designed according to the structure information of target enzyme binding sites. Thirdly, the building blocks are designed to contain different structures to cover different geometrical and functional spaces for chemical diversity generation.

Usually, the design of the building blocks can be based on an existing inhibitor of the target enzyme[18], by analysis of the target enzyme structure[19], or by molecular modelling[20]. For example, a dynamic combinatorial library was designed based on ketones and amines as

building blocks to discover potent inhibitors towards neuraminidase, a key enzyme in the influenza virus activity.[18] Diamine was used as the scaffold for the dynamic library as it is structurally similar to known neuraminidase inhibitors, like zanamivir[21]. A series of ketones with different structures was used for the imines formation (**Figure 47**). The imines were then reduced to amine products because of their instability and the composition of DCL was monitored by LC/MS. The result illustrated that the distribution of imines library members altered due to the addition of enzyme. The amplified amines were then tested against target enzyme to verify their inhibitory activity.

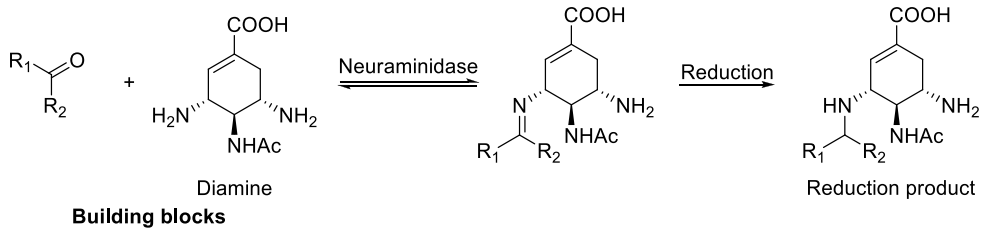


Figure 47 Design of imine DCL of ketones and amine building blocks (modified from reference[18])

It was reported that a DCL based on imine formation was generated for the development of inhibitors towards GaTs (including α-1,3-Galactosyltransferase (α-1,3-GalTs) and β-1,4-GalTs).[22, 23] Uridine aldehyde, an analogue of uridine, was used as an anchor building block because the nucleoside moiety enables to locate DCL members towards active sites of GalTs.[24] A series of amine was selected as surrogates of the pyrophosphate and galactose moieties, constructing imine DCL. The most amplified binder was identified in the presence of enzyme and the corresponding amide product of this amine library member was assessed to be an inhibitor towards enzyme, with IC_{50} of 0.4 mM. (**Figure 48**) This successful development of GalT inhibitors *via* a DCC approach directly inspired our own attempts to

identify suitable pyrophosphate and sugar replacements in our series of 5-substituted nucleoside GalT inhibitors *via* DCC experiments (see below).

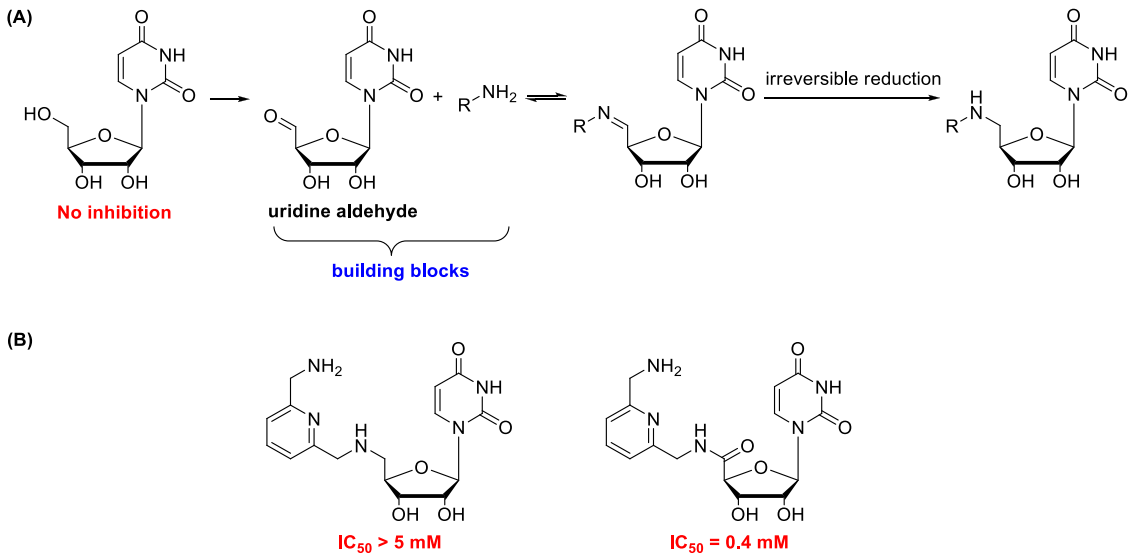


Figure 48 (A) Design of DCL of uridinal and amine building blocks. (B) Structures of corresponding amine product containing the best binder and its amide analogue.

In another example, a protein-directed DCC experiment was applied to develop inhibitors for carbonic anhydrase (CA).[25] X-ray crystallographic data of target enzyme has shown that a Zn(II) ion is situated in the active site, being coordinated by Histidine residues. Sulphonamides can act as recognition element to bind with CA by adding to the metal coordination sphere.[26] According to this structure information, a benzhydrazide derivative containing $ArNH_2SO_2$ group was employed as building blocks. A series of aromatic aldehydes were selected as well to introduce 'tails' onto the benzhydrazide derivative which might complement the interaction between the sulphonamide moiety and enzyme (**Figure 49**). The effective building blocks in the DCL were then distinguished by using mass spectrometry.

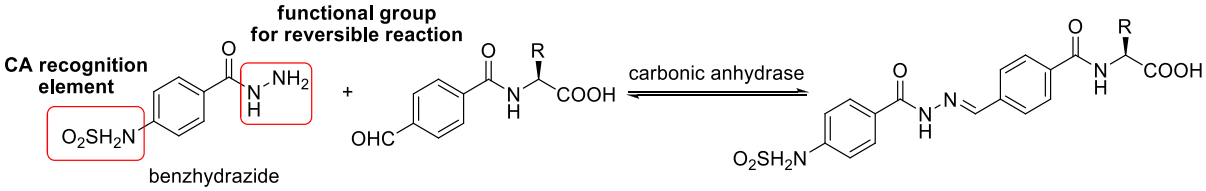


Figure 49 Design of DCL of aromatic sulphonamide and aromatic aldehydes (modified from reference[25])

3.1.2.3. Reversible reactions and reaction conditions for protein-directed DCC

To date, various dynamic reversible reactions have been used in protein-directed DCC. They can be classified into three main types (**Table 10**):

(i) Addition-elimination reactions of carbonyl groups, such as imine formation[14,18, 27] and hydrazone formation[19, 28].

(ii) Thiol exchange reactions, such as thiol-disulphide exchange[13], thiol-enone reaction[19] and hemithioacetal reaction[29, 30].

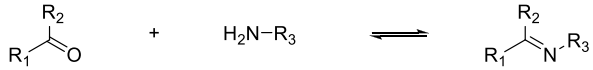
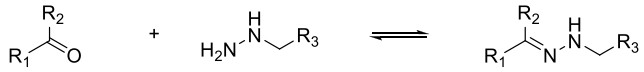
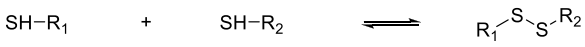
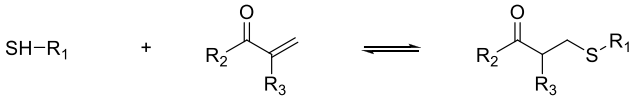
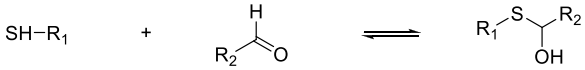
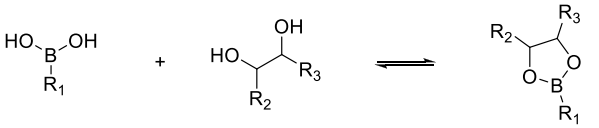
(iii) Boronate ester formation[15, 31].

In addition to these reactions,, some organic-based reactions, like alkene cross metathesis[32] and oxime formation[33] also have been employed in DCC experiments.

The selection of an appropriate reversible reaction is crucial for the design of protein-directed DCC. The reversible reactions need to fulfil several requirements.[3, 17] The reversible reaction can be carried out in aqueous media, which makes them biocompatible. The reversible reaction chosen for DCL should equilibrate on a reasonable timescale, from minutes to days, under mild conditions, at a certain pH and temperature, where target proteins are active and molecule recognition would not be interfered. The reversible reactions can be carried out in

the presence or absence of enzyme. It needs to guarantee the solubility of all the library members at equilibrium that the DCL would not be biased due to precipitation of some library members. Also, the reversible reactions can be carried out chemoselective to avoid cross-reacting with functional groups of building blocks.

Table 10 Reversible reactions used in protein-directed DCC

Reaction	Scheme
imine formation	
Hydrazone formation	
Thiol-disulphide exchange	
Thiol enone exchange	
Hemithioacetal reaction	
boronate ester reaction	

Imine formation

Imine formation was one of the first reversible reactions employed for protein-directed DCC.[7] In DCL, imine libraries are generated by the condensation between aldehydes or ketones/amines at mildly acidic and room temperature conditions. Because of the instability of some imines, reducing reagents could be introduced to reduce instable imines into amine.

The kinetic information of these libraries was analysed by the detection of stable amine analogues.[14, 22] (**Figure 50**)

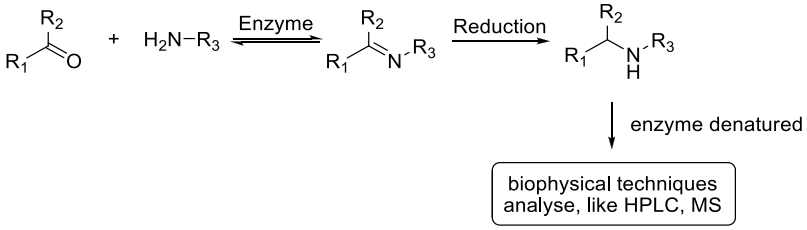


Figure 50 The imine library generated *via* ketone and amides and the imines were reduced for analysis

A DCL was generated by using the imine interconversion for the identification of inhibitors of human carbonic anhydrase (hCA II).[34] (**Figure 51**) Different aromatic aldehydes and amines were selected as building blocks for probing the active enzyme site. In this DCL, 20 imines were produced in equilibrium and the reducing reagent $NaBH_3CN$ was applied to stop the reversible reaction, generating corresponding amines. The HPLC analysis results showed that the distribution of the mixtures was altered by addition of the target enzyme and the amplified members were identified. All possible amines were synthesized and inhibition constants (*K*i) were determined against hCA II.

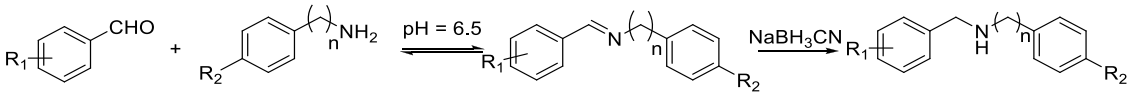


Figure 51 Reversible formation of DCLs based on imine formation for the development of hCA II inhibitors (modified from reference[34])

Hydrazone formation

The reversible formation of hydrazones was first used by the Sanders group for DCL synthesis.[35] The hydrazone exchange reaction is carried out between aldehydes/ketones and hydrazides. Compared with imine formations which react at neutral to acidic condition, hydrazones can be thermodynamically stable at lower pH condition, ranging from 2.5 to 6.0, as they are favoured by the protonation of the aldehyde and the nitrogen atom of the C=N bond of the hydrazone.[1] Under acidic or high temperatures conditions, the exchange reactions become significant, which gets slower as pH increases. This reaction is chemoselective and thus compatible with various functional group, highly suitable for DCLs. DCLs based on hydrazone exchange have been reported by using nucleic acids, carbohydrate or peptides as building blocks.[36, 37]

In 2003, the inhibitors of cyclin-dependant kinase (CDK2) which is involved in various human cancers was developed by generating DCL of hydrazones.[38] A range of hydrazines and isatins were selected as building blocks and the hydrazone formations were carried out in neutral condition. (**Figure 52**) As they tend to be kinetically inert under neutral conditions, the reaction equilibrated in relative long timescale, 48h, at room temperature. The reactions were carried out in the presence of protein and the amplified hydrzones were identified. Biochemical evaluation of these hydrazones showed that all of them exhibited potential inhibitory activity, with IC_{50} values of around 30 nM.

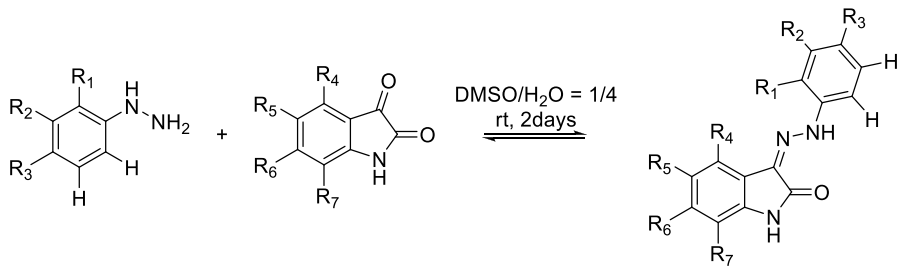


Figure 52 DCL of oxindole compounds synthesized by reaction of hydrazine and isatins for the development of CDK2 inhibitors. (modified from reference[38])

Disulfide bond formation

Disulfide bond formation requires deprotonating thiol, which is pH-dependent and it can be quenched by decreasing pH. Neutral condition (pH 7-9) is appropriate for the generation of disulfide formation. Due to the neutral reaction condition for disulfide exchange reaction, this reversible reaction is also one of most common reaction applied for protein-directed DCC. The first example of disulfide-based DCC was introduced by Lehn.[13] Then the disulfide-based DCC has been used for development of inhibitors towards biological targets.[39, 40]

Disulfide-based DCC was employed for targeting adenosine-binding sites on *Mycobacterium* tuberculosis pantothenate synthetase, a potential virulence factor for M. tuberculosis that is required for pantothenate (vitamin B5) biosynthesis.[41] 5'-deoxy-5'-thioadenosine was chosen as an anchor to the adenosine binding site of enzyme and a set of thiols was selected as building blocks as well. (**Figure 53**) Upon addition of the protein, a benzyl disulfide derivative was amplified and identified on library analysis by HPLC. As predicted by a molecular docking study, the benzyl substituents can be accommodated in the hydrophobic substrate-binding pocket. This hypothesis was confirmed by solving the co-crystal structure of the target enzyme complex with the hit disulfide compound.[41]

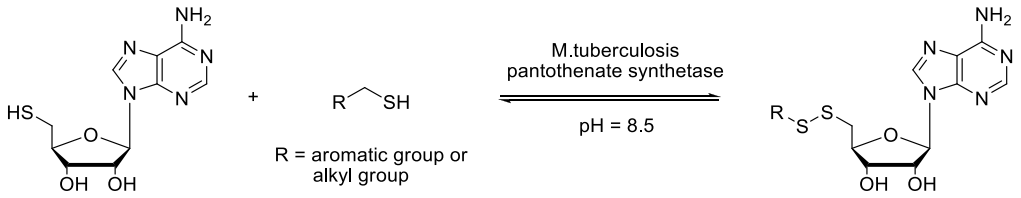


Figure 53 Generation of disulfide-based DCL for templating by M.tuberculosis pantothenate synthetase

As the disulfide library often requires long incubation time to reach equilibrium, from several days to two weeks,[42, 43] improvements were developed to reduce the incubation time. For example, in the protein-directed DCC work based on the disulphide reversible reaction, the experiment was carried out by using glutathione redox buffer at pH 8.5 under inert atmosphere.[41] Because the glutathione redox buffer could promote disulphide exchange, the DCL reached equilibrium in less than 24h.

Thiol-enone exchange

The thiol-enone exchange reaction is performed by nucleophillic attack of thiol towards enone (α,β-unsaturated ketone). This reaction is suitable for protein-templated DCC because this exchange reaction can reach equilibrium fast (from hours to days) and take place in aqueous media.[40, 44, 45]

The reversible reaction between thiols and enones was used in protein-directed DCC towards the target enzyme glutathione S-transferase (GST).[45] GST and its analogues as the thiol building blocks and ethacrynic acid as the enone were chosen because it was a known inhibitor of enzyme (**Figure 54**). The thiol and the enone afforded thioesters, which reached equilibrium in 1h. Upon the addition of target enzyme, the reactions re-equilibrated, which indicated that the enzyme could facilitate the addition between thiols and enone. In order to confirm the binding affinity between enzyme and the amplified products, these compounds were synthesized and subjected to the standard GST assay for assessment. Inhibition results showed IC_{50} value of 0.32 μM.

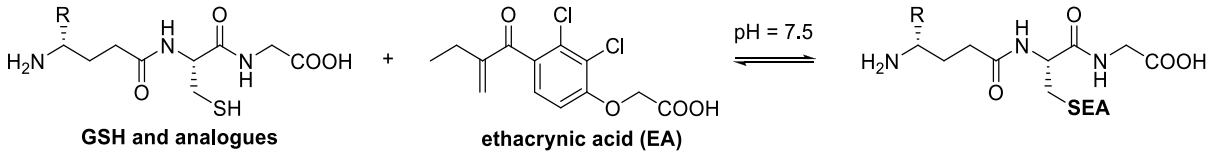


Figure 54 Generation of thiol-enone based DCL via GSH and ethacrynic acid for templating by GST

Hemithioacetal reaction

The hemithioacetal reaction is performed between thiol and aldehyde. The hemithioacetal formation can be thermodynamically stable at neutral pH condition and it can reach equilibrium in a short timescale. For example, the hemithioacetal formation was applied for protein-directed DCC and β-galactosidase was selected as the target protein for the proof-of concept study.[30] Based on the structures of known inhibitors of target enzyme, *o*-nitrophenyl-β-galactopyranoside and isopropyl-1-thio-β-D-galactopyranoside, a set of thiols and aldehydes were selected as building blocks, constructing hemithioacetal DCL (**Figure 55**). Because the formation and dissociation of hemithioacetal are fast compared with NMR relaxation time at pH = 7.2, Saturation-Transfer Difference NMR (STD-NMR) was applied to identify the binders of enzyme. It was shown that hemithioacetal products could be selected as binders by target enzyme, which was confirmed in the inhibition assay.[30]

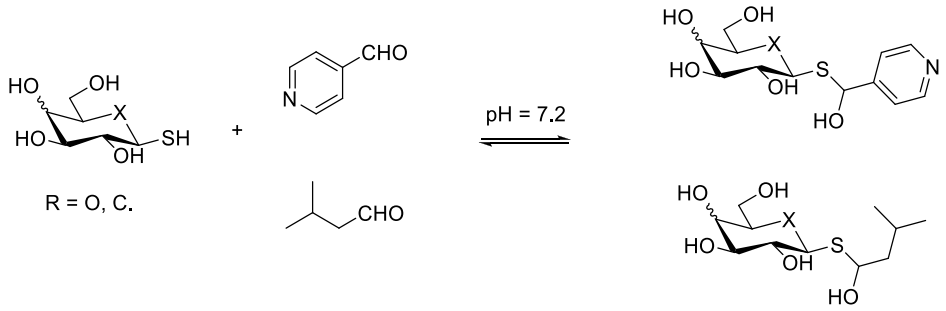


Figure 55 Generation of hemithioacetal-based dynamic combinatorial library of thiols and aldehydes for the proof-of-concept study

Boronate ester formation

Boronate ester formation is performed between boronic acids and diols. The reversible reactions reach their most dymanic at the pH which is similar to the pK_a of the boronic acid.[3] Typically, arylboronic acid's pK_a are around 8, which makes the reversible reaction suitable for neutral condition. The pK_a value of alkylboronic acid are higher (10 -12) so that basic reaction condition is suitable for reversible reaction with alkylboronic acid. Therefore, it is essential to measure the pK_a of the boronic acid and choose an appropriate pH which is suitable for both boronic acids as well as the target enzyme.[15]

DCC experiments by using boronic ester reversible reaction were carried out by Leung to enhance the inhibition of serine protease α-chymotrypsin.[15] Based on known serine-protease inhibitors were selected as model boronic acid fragments, and fructose were selected as the alcohol fragments for use in DCC (**Figure 56**). ^{11}B-NMR and ^{1}H-water-ligand observed via gradient spectroscopy (^{1}H-waterLOGSY) were used to monitor the ternary complexes of enzyme, boronic acids and sugars. Given that the pK_a of most boronic acid is around 7-9, the DCC experiments were conducted at pH 5.8 to make sure that the free boronic acid is present in its sp^2 hybridized form to react with diol in the following step. In the presence of enzyme, the favourable combination of boronic acids and diols were identified by NMR analysis.[15]

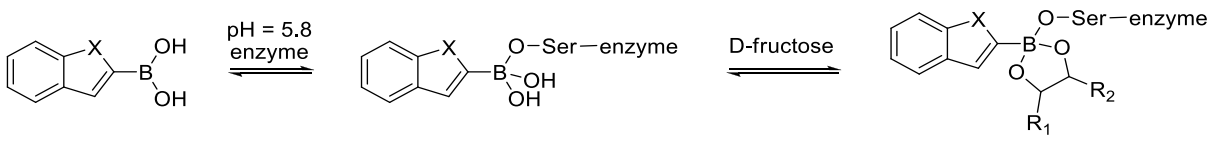


Figure 56 Generation of boronate ester based DCL via boronic acid fragment and fructose, templating by serine-protease

3.1.2.4. Analytical techniques

Because of the increasing interest in DCC for the development of potent inhibitors and drug discovery, a lot of attention has been paid to the development of suitable analytical tools for the identification of binders from DCLs. The techniques need to be sensitive enough to detect changes in equilibrium and concentrations, and they need to be able to detect protein binders from a mixture of reversibly reacting products.[3, 46] Until now, three main types of analytical techniques have been used, namely high pressure liquid chromatography (HPLC)[18, 27, 47], mass spectrometry[19, 48, 49] and NMR spectroscopy[15, 29, 31, 41].

HPLC is routinely used for the isolation and identification of small molecules. For DCLs, HPLC is particularly useful as it can screen the binders and quantification of the relative concentrations from the mixtures of DCLs. The comparison of HPLC chromatograms of the DCLs mixtures under different thermodynamic conditions can help for the identification of amplified binders.[27, 50, 51] However, one of the disadvantages of HPLC for DCLs is that the unstable DCL members may dissociate when interacting with the mobile and the stationary phases of HPLC in the process of separation. In order to avoid this problem, the reversible reactions of DCC are often 'quenched' prior to HPLC analysis. For example, reductive reagents, like $NaBH_3CN$, were used to imine exchange reaction to produce corresponding stable amine to frozen the reaction prior to HPLC analysis.[14, 47] pH modulation is carried out for thiol disulphide exchange reaction before analysis because the deprotonation of thiols is pH-dependant and the decreasing pH could freeze the exchange reaction.[40] Also, by increasing pH of a mixture of hydrazones could slow down the reversible reaction as they are much more stable at neutral or basic pH condition.[37] The pH of exchange reactions can be adjusted by the mobile phase of HPLC. For example, a DCL based on the imine reversible

reaction was developed to identify inhibitors of egg white lysozyme.[27] The sample of DCLs mixture was applied to the HPLC analysis in which mobile phase contained TFA (0.1%). It effectively decreased the pH of DCLs mixtures and 'quenched' the equilibrating mixture. In catalysed reversible reactions, the removal or inactivation of the catalyst will transform the DCLs into a stable library for subsequent analysis.[52]

NMR spectroscopy is a non-destructive analytical technique, which can be performed directly in solution. The NMR technique can provide structural information of compounds and It can be employed for quantitative analysis as well. For DCLs, NMR tools permit the direct analysis of the complete mixture under equilibration conditions. For example, the ^{1}H-NMR experiments are a useful tool for the quantitative measurement and for the optimization of DCLs applications.[53] ^{13}C-NMR can also be applied for the analysis of DCL.[54] It is possible that ^{1}H and ^{13}C-NMR signals of individual compound become overlap with the increasing number of DCLs, which interferes with the detection and quantitation of DCLs. In order to enhance the sensitivity of NMR tools for detecting, a common approach is the insertion of ^{13}C-isotopic labels at key positions of the molecular structure of DCLs.[54, 55] The ^{13}C-labeling of the imine carbon of DCL members provides a single characteristic resonance for each compounds in the ^{13}C-NMR spectrum which could increase its sensitivity close to that of a ^{1}H-NMR spectrum, which permit the increasing complexity of DCLs.[55] Also, the ^{13}C-isotope would not affect the thermodynamic stability of exchange reaction. ^{19}F-NMR is applied as an alternative approach for the determining the thiol-enone exchange equilibrium constant, where the DCLs members, thiol-enones, are appropriately labelled.[56] In the boronate ester formation DCL for α-chymotrypsin inhibitors development, ^{11}B-NMR and ^{1}H water LOGSY were used as tools to observer ternary complexes of boronic acids, diol and enzyme for analysis.[15] Since the ^{11}B-NMR is relatively insensitive, the concentration of target enzyme was increased to 2 mM (ratio: enzyme/boronic acid = 1/1). Also, the nuclear overhauser

effect (NOE)-based NMR techniques such as STD[57] and water ligand-observed via gradient spectroscopy (waterLOGSY)[15, 41] are applied to detect binders of target enzymes in DCLs. For example, a hemithioacetal formation was employed for the generation of DCL for β-galactosidase and the binding performances of DCL members were detected by STD NMR technique.[30] As the hemithioacetal formation and dissociation processes are much faster than the normal NMR relaxation time scale under the enzyme-suitable condition, their resonances were broadened and difficult to be observed. The STD-NMR only focuses on the interaction between ligands and target protein. The magnetization, which is transferred from enzyme protein protons to those of bound ligands, can be detected after dissociation from the enzyme by STD-NMR. In another example, boronate ester formation DCL was applied for the development of α-chymotrypsin inhibitors by using ^{11}B-NMR and ^{1}H water LOGSY as the analytical method of choice.[15] A DCL containing 5'-deoxy-5'-thioadenosine and thiols as building blocks was developed, with an adenosine-binding enzyme as template.[41] A ^{1}H-NMR waterLOGSY experiment was employed for detecting the binding between protein and thioadenosine building blocks. The purine protons from thioadenosine, which were distinct from other signals on spectrum, can be detected strongly positive upon the protein-ligand binding.

Mass spectrometry is also a commonly used technique to identify molecules because it can directly confirm the presence of a particular molecule. In DCLs, electrospray ionization mass spectrometry (ESI-MS) is often used to detect the produce of DCL members.[58-60] The ESI-MS screening has permitted concurrent identification of all ligands of enzymes through direct analysis. The approach distinguishes the effective combination of building blocks in the DCL by specific detection of the target protein-ligand complexes. For example, a thiol-based DCL was combined with a ESI-MS approach to identify inhibitors of the BcII metallo-β-lactamase (MBL).[60] As the compound *meso*-2,3-dimercaptosuccinic acid was indicated to be able to

bind with the complex of MBL and Zn(II) ion and their peaks could be observed by ESI-MS, this compound, used as an anchor building block, and a set of thiols were applied to DCL in the presence of enzyme. After a short time of aerial exposure, peaks corresponding to BcII-Zn-disulfide complexes could be observed by ESI-MS selectively. The selected disulfides were confirmed to act as a potent inhibitor towards MBL (Ki = 125 μM). A disadvantage of this technique is that its use for quantification is not straightforward. In order to solve this problem, the combination of liquid chromatography and MS (LC/MS) has been applied as a very powerful approach for DCLs analysis, which allows the separation and identification of the DCL mixtures.[61] The separation of the mixture by HPLC can give information about the relative amount of compounds present in the library and the following analysis of MS allows the identification of each species.

Other biophysical techniques also have been developed for the screening of DCL binders. For example, X-ray crystallography has also been applied to DCL experiments. In 2003, X-ray was used to detect binders of cyclin-dependant kinase 2 (CDK-2) from a DCL by exposing protein crystals to the library.[38] The building blocks, hydrazines and isatins, was exposed to CDK-2 crystals and one of the DCL members was observed to locate in the hydrophobic region of enzyme active site and it clearly indicated that X-ray crystallography could act as a useful tool to detect enzyme binders. It is obviously that small size DCLs are more easily to be employed as the complexity of mixtures of large size DCLs increases. The number of building blocks is increasing and it will result in the larger amount of equilibrating compounds. In this condition, false positive results might be produced. In order to solve the problem, a DNA-encoding technology was applied with protein-directed DCC to identify the amplified compounds and it allowed efficient screening of potentially large amount of compounds.[62, 63] In the case that DCLs are composed of aromatic and heterocyclic

compounds, the use of NMR technique may lead to false positive results and other techniques, like UV spectroscopy, were applied to verify the results.[29]

3.2. Objectives

Based on the results of our previous experiment, uridine derivatives were generated by the removal of pyrophosphate and galactose moieties and the introduction of 5-position substituents. These 5-substitunt nucleoside derivatives were identified as nucleoside-based β-1,4-GalT inhibitors. The IC_{50} of the nucleoside derivatives were 200 to 300 μM, twenty times weaker than the sugar nucleotide inhibitor, 5-FT UDP-Gal. It is speculated that the drastic reduction of activity was due to the absence of pyrophosphate and galactose moieties. A dynamic combinatorial chemistry strategy was applied to identify appropriate replacements of pyrophosphate and sugar moieties (**Figure 57**) in order to overcome the loss of activity and develop potent β-1,4-GalT inhibitors. It was speculated that the pyrophosphate and galactose mimics that were screened from the DCL could exhibit binding affinity toward the enzyme, thus facilitating the inhibitory activity of corresponding nucleoside derivatives.

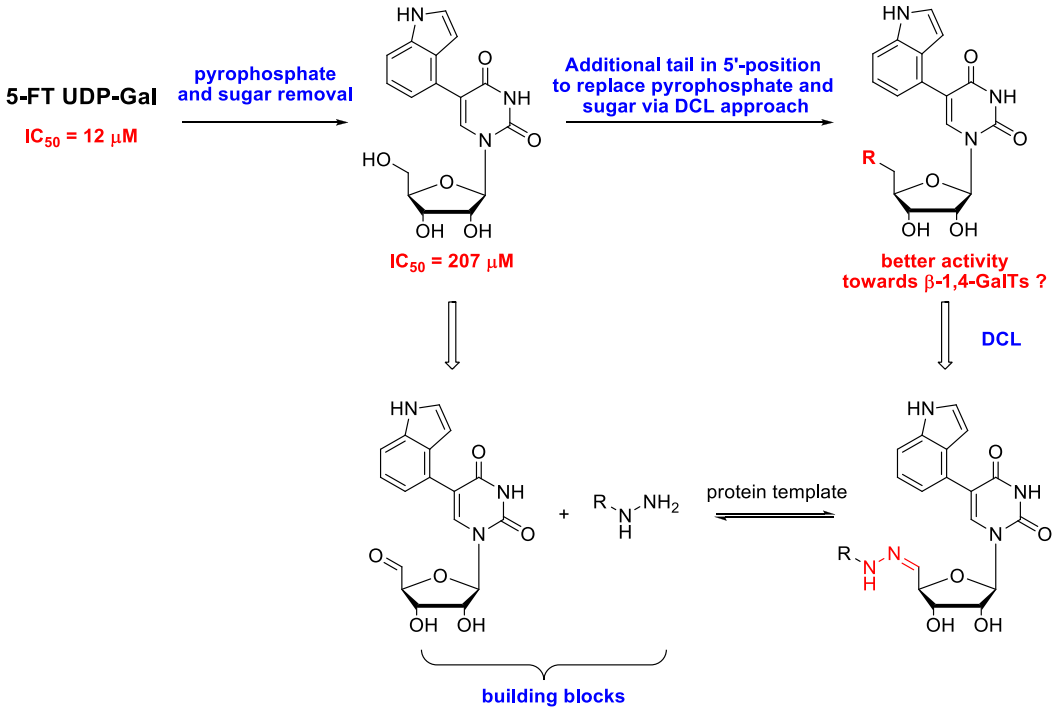


Figure 57 Generation of a hydrazone DCL from 5-indole uridine aldehyde and hydrazides/hydrazones

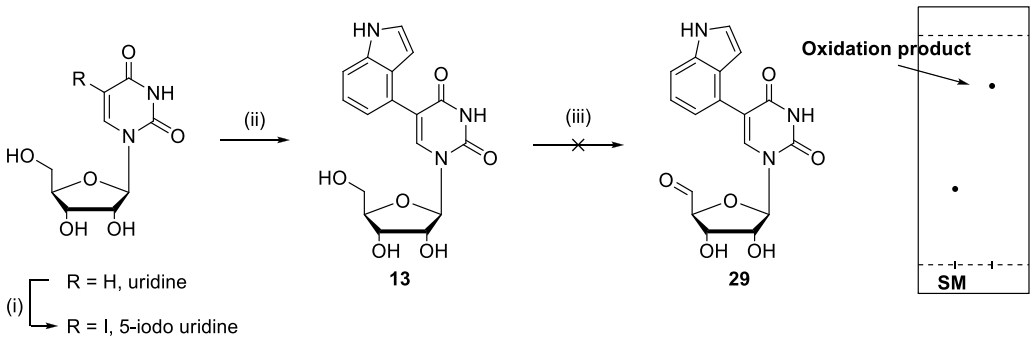
3.3. Results and discussion

The synthesis of the 5-indole uridine aldehyde is initially introduced in this section before describing the development of the DCC experiment and HPLC-based analytical method. This is followed by a synthesis and activity assessment of the active DCL members.

3.3.1. Synthesis of building blocks (5-indole uridine aldehyde)

In order to start the DCC experiment for inhibitors development, the synthesis of building block, 5-indole uridine aldehyde **29**, was carried out. It was speculated that, since 5-indole uridine contains three unprotected hydroxyl groups, one primary hydroxyl group and two

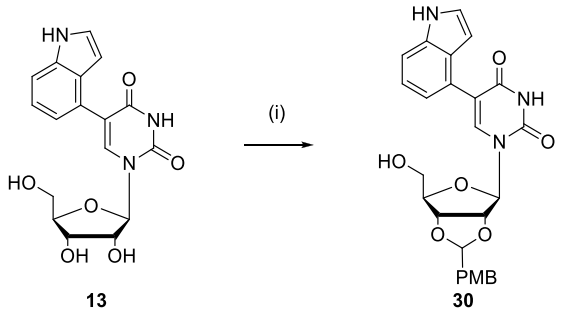
secondary hydroxyl groups, the selective oxidation of the primary hydroxyl group might be achieved without protecting the secondary groups. Although there was no precedent for the selective oxidation of nucleoside primary groups, it was decided to try it due to the simple, one-step synthesis. Therefore, the direct oxidation of the primary hydroxyl group was tested after obtaining 5-indole uridine **13** *via* iodination of uridine and subsequent Suzuki reaction under MW conditions. (**Scheme 9**) Since the solubility of starting material **13** was low in DCM, DMSO was used as the solvent to facilitate its dissolution. The reaction was carried out in DCC oxidation condition, and after stirred in room temperature for 3 hour, TLC was applied to monitor the reaction. Based on the TLC, apart from some starting material that was left behind, a new spot with a much higher R_f value appeared on the plate. Crude mixture was used for ^{1}H-NMR. The 2'-, 3'-, 5'-position protons did not appear in the NMR spectrum. Based on the results of ^{1}H-NMR and TLC, it was hypothesised that all the hydroxyl groups had been oxidized.



Scheme 9 Synthesis of protected 5-indole uridine. Reagents and conditions: (i) I_2, HNO_3 (2M), $CHCl_3$, reflux for 6h, 70%. (ii) 4-indole boronic acid, Cs_2CO_3, TPPTs, $PdNa_2Cl_4$, H_2O, MW 120 ºC, 30 min, 45 %. (iii) Trifluoroacetic acid, DCC, pyridine, DMSO, 3h, product with all OH groups oxidized.

Therefore, it was necessary to carry out the protection step of diol prior to the oxidation of primary hydroxyl group. It has been reported that commercial uridine reacted with *p*-methoxybenzaldehyde to produce the 4-methoxylbenzyl ether (PMB)-protected compound in

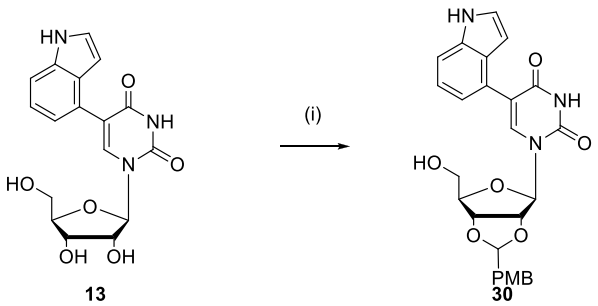
high yield[64] and the PMB protecting group can be easily cleaved off later under very mild conditions using 5% trifluoroacetic acid (TFA) in DCM. Therefore, compound **13** was applied to the protected reaction. (**Scheme 10**)



13 30

Scheme 10 Synthesis of PMB protected 5-indole uridine. Reagents and conditions: (i) *p*-methoxybenzaldehyde, $ZnCl_2$, dry THF, rt, 2 days.

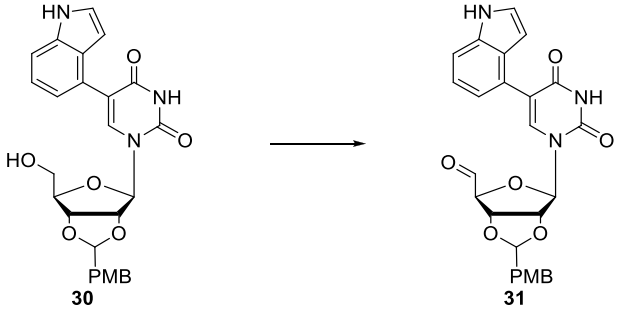
The reaction was carried out using *p*-methoxybenzaldehyde (2 equiv.), $ZnCl_2$ (1 equiv.) in anhydrous THF at room temperature. However, only starting material was obtained after 48 hours. The used batch of $ZnCl_2$ was not powder but like large pieces. It was speculated that the $ZnCl_2$ was not anhydrous due to its high hygroscopicity, which may have interfered with the protective reaction. Commercial uridine was subjected to this reaction in the presence of a new batch of $ZnCl_2$ in order to verify this hypothesis, and a satisfactory 86 % yield of desired product was obtained. The compound **30** was prepared in the same way, but only a poor purified yield of 27% was obtained. The protocol employing *p*-methoxybenzaldehyde dimethyl acetal (2 equiv.), *p*-TSA (5 mol%) in dry DMF solution at 70 ^{o}C overnight was applied in an attempt to improve the efficiency of this protective reaction and this produced the desired compound **30** in a yield of 68 %. (**Scheme 11**)



Scheme 11 Synthesis of compound **30**. Reagents and conditions: (i) *p*-methoxybenzaldehyde dimethyl acetal, p-TSA, dry DMF, 70 ºC, overnight, 68%.

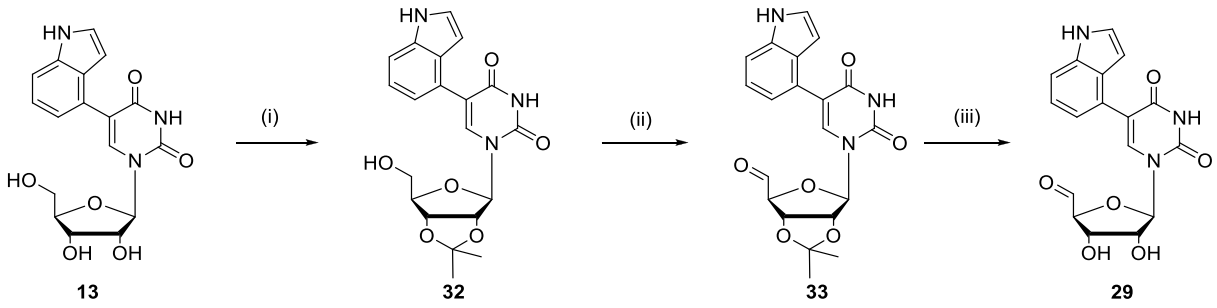
With the protected precursor **30** in hand, the synthesis of compound **31**, PMB-protected 5-indole uridine aldehyde, was started by the oxidation of the DCC. (**Table 11**, **entry 1**) The reaction was carried out using TFA, DCC, and pyridine in a DMSO solution at room temperature for 6 hours. The TLC results showed that the reaction mixtures were quite complex and no major spot was detected with strong UV absorbance. Then, the Dess-Martin oxidation was investigated using Dess-Martin reagent (3 equiv.) in DCM, as shown in **entry 2**, the result was similar as **entry 1**.

Table 11 Oxidation of PMB protected 5-indole uridine aldehyde **31**



Entry	Reactions	Results
1	DCC, TFA, pyridine, DCM, rt, 6h	NO product
2	Dess-martin oxidation	NO product

Since the oxidation step had failed, a new strategy was adopted by using a different protective group (**Scheme 12**). It has been reported that 2,2-dimethoxy propane (DMP) can be used to protect diols in the form of isopropylidene under acid catalysed conditions.[65] Therefore, DMP protection was carried out by using DMP (3 equiv.), *p*-TSA (5 mol%) in a THF solution at 60 ^{o}C for 3 h and the desired product **32** was obtained in a yield of 69% after purification. Dess-martin oxidation was performed by subjecting compound **32** to the condition of Dess-Martin reagent in DCM solution at rt for 5h. The desired product **33** was obtained in a yield of 47%. Then the DMP protective group of compound **33** could be cleaved off later using 50% trifluoroacetic acid (TFA) in DCM for 30 min. Following the removal of the excess TFA solution by evaporation, the desired aldehyde **29** was obtained in 80 % yield without further purification.



Scheme 12 Synthesis of 5-indole uridine aldehyde. Reagents and conditions: (i) 2,2-dimethylpropane, *p*-TSA, TFA, 60 ^{o}C, 3h, 69%. (ii) Dess-martin reagent, DCM, 5h, 47%. (iii) 50 % TFA, 30 min, rt, 80%.

3.3.2. DCL screening experiment

With the desired nucleoside building block **29** in hand, the DCC experiment was carried out. A DCL was generated based on a hydrazone formation using compound **29** and

hydrazines/hydrazides (**H-1** to **H-13**) as building blocks. A series of hydrazines/hydrazides with different structures was selected to establish library diversity and they were all commercially available. (**Figure 58**) It has been reported that tunicamycin acts as GlcNAc phosphotransferase inhibitor by utilizing a monosaccharide to mimic the pyrophosphate-Mn^{2+} six-member ring complex of transition state. Glucose residues were also investigated as possible pyrophosphate-metal complex mimics for the GalT inhibitor development. Thus, **H-1** and **H-4** containing six-member rings were chosen in order to mimic the six-member ring complex formed in the transition state. Aliphatic hydrazines/hydrazides containing functional groups, like **H-3**, **H-11** and **H-5**, etc., were selected due to the possibility of binding between enzyme active site and their functional groups. The indole moiety could help the binding towards β-1,4-GalTs, as was proven in the development of our nucleotide inhibitors; therefore, the hydrazide that contains indole moiety, **H-13**, was selected as the building block.

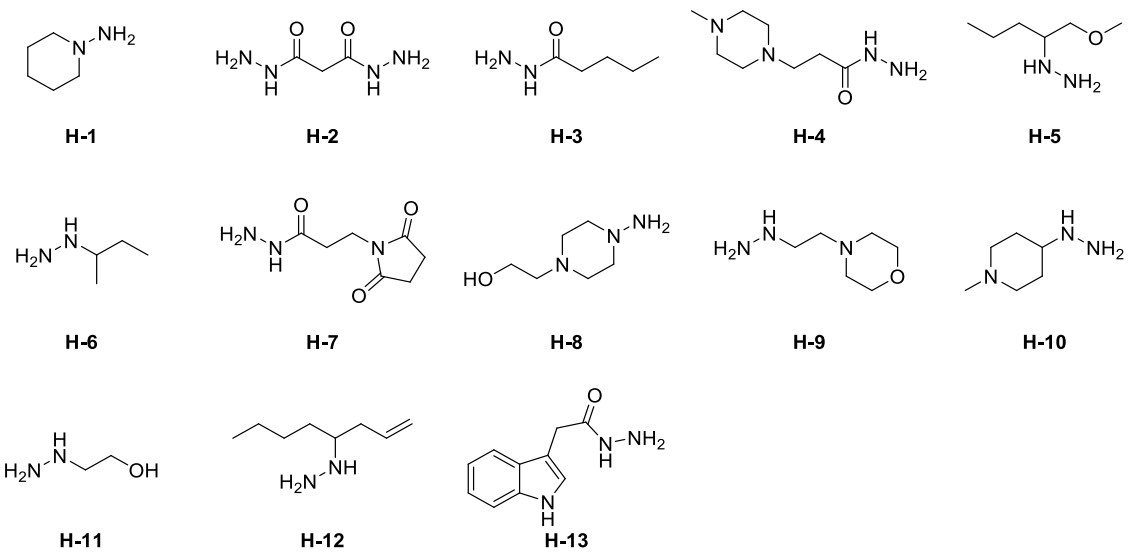


Figure 58 Library of hydrazine/hydrazide building blocks

Since the target enzyme, β-1,4-GalTs, was the same as the ones in the literature, it was decided to start the experiment under the conditions as in the literature,[22] according to which

the reversible reaction reached an equilibrium after 13 days at room temperature. Although a different reversible reaction was used in our DCC experiment, it was possible that it would still take a long time to achieve equilibrium. Therefore, in view of the potentially long incubation time, we decided to first test the activity of the enzyme under the conditions of the DCC experiment, in order to avoid false screening results. The enzyme was incubated in the DCC experiment buffer (MOPS buffer, 50 mM, pH 7.4, $MnCl_2$ 20 mM) in a water bath at 25 ^oC. The enzyme activity was monitored by measuring the percentage substrate UDP-Gal turnover in phosphatase-coupled glycosyltransferase assay every other day (14 days in total), (**Figure 59**). The turnover% of the natural donor UDP-Gal was plotted against the incubation time (days) and it was suggested that the enzyme activity decreased as the turnover% dropped. The turnover% was still around 17% after 14 days of incubation, which verified that the enzyme could have a catalytic effect on the reversible reaction during the incubation process.

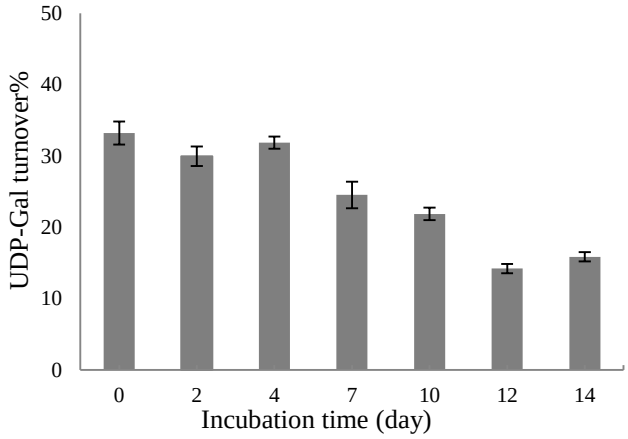


Figure 59 GTs activity assay by measuring the conversion of the UDP-Gal donor at different incubation time (with ‘’turnover’’ defined as the percentage of UDP-Gal that is consumed during the GT reaction, relative to the total amount of UDP-Gal in solution)

Since the reversible reaction may take a long time to reach equilibrium, the library products formed may not be stable. Due to the instability of these hydrazone products, the reducing reagent $NaBH_3CN$ was introduced to reduce the hydrazone products into stable products,

which were irreversible. Then the distribution of reduced products could be analysed by HPLC.

The concentration of components in the reaction was the key points of the DCC experiment. For example, the concentration of building blocks can influence the rate of reaction. It is crucial that the reversible reaction can reach equilibrium in a reasonable timescale. Also, a sufficient amount of enzyme is required to re-equilibrate the reversible reaction and the changes in the distribution of library members are detectable for the analytical technique. An excessive amount of reducing reagents was used to facilitate the reductive step, resulting in a stable product for analysis.[66] It was reported in the literature that no products could be detected by HPLC-UV or HPLC-MS after 13 days incubation when the initial concentration of building block was 80 μM.[22] Thus, the concentration of building blocks was increased to 300 μM, in order to push the equilibrating mixture towards product formation for analysis. Therefore, our initial attempt of the DCC experiment was tested utilising compound **29** and piperidin-1-amine **H-1** as building blocks with a concentration of 300 μM in a MOPS buffer (50 mM, pH 7.4, $MnCl_2$ 20 mM) as a model reaction. The 5 mM of the reductive reagent $NaBH_3CN$ was much in excess of the building blocks, and the enzyme was introduced with a concentration of 0.69 mg/mL. The total volume was 300 μL. A control experiment containing identical reagents except the enzyme was performed in order to compare the re-equilibrium effect from enzyme template. The reactions were incubated at 25 ^{o}C and monitored every day by HPLC.

Since the aldehyde building block **29** and formed product are UV-active and their UV-absorbance are both derived from the nucleoside and indole moieties, the absorbance spectrum of the aldehyde derivative **29** as well as 5-indole uridine was measured to select the wavelength of HPLC detection. 25 μM aldehyde **29** and 5-indole uridine in methanol solution

were measured separately. (**Figure 60**) The maximum absorbance of the 5-indole uridine, as well as the aldehyde derivative, was around 262 nm and their UV absorbance at 254 nm was found to reach maximum intensity of 80%. Therefore, absorbance at 254 nm was selected for the HPLC analysis.

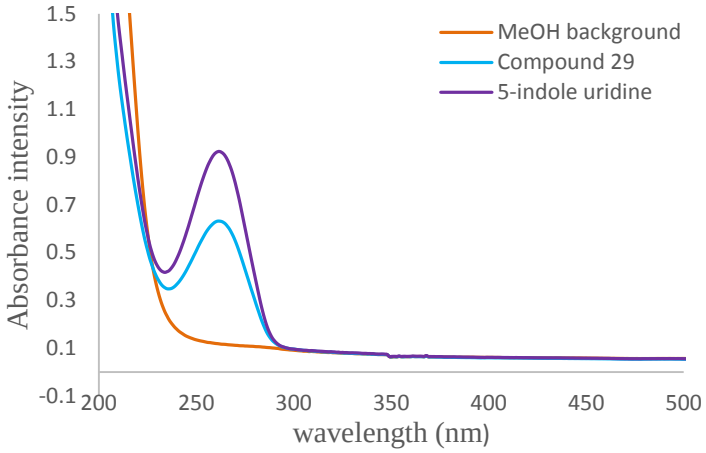


Figure 60 UV absorbance measurement of 5-indole uridine and its aldehyde derivative **29**. Maximum absorbance at 262 nm and 80% maximum absorbance intensity at 254 nm.

A calibration curve was established to measure the relationship between the sample volume and UV absorbance intensity. Different volumes of aldehyde **29** solution (300 μM) were applied to the HPLC equipment. The volume of **29** solution was plotted against the UV-absorbance intensity to obtain the calibration curve. (**Figure 61**) According to the calibration result (the equation of calibration), the sample of which the volume was lower than 50 μL was un-detectable for the HPLC. Thus, the injected amount of reaction mixture solution was raised to 100 μL.

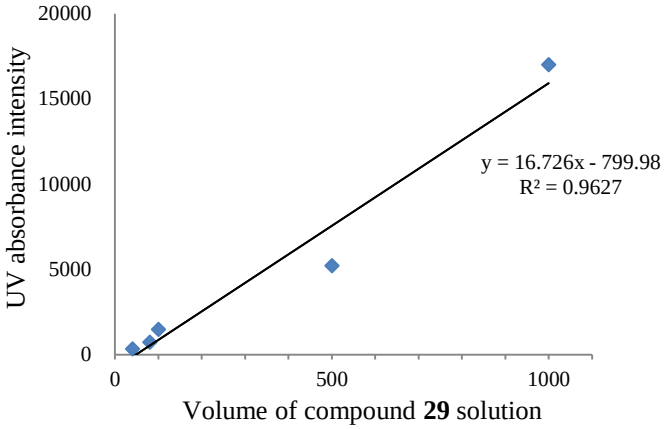


Figure 61 The calibration curve for the measurement of relationship between compound volume and UV-absorbance intensity. Different volume of **29** solution (10, 40, 80, 100, 500, 1000 μL) was injected in HPLC and detected on 254 nm.

In the reversible reaction, the mixtures of **29** and hydrazines/hydrazides were equilibrated, forming hydrazone products, which were then reduced to stable compounds. The equilibrium of the reversible reaction in the presence of enzyme may induce a change in the distribution of the reduced products, which can be analysed by HPLC. Thus, a robust method was developed to separate the **29** and reduced products which were UV-active. An HPLC system with detection set at 254 nm coupled with a reverse phase column (150 × 4.6 mm) with a flow rate of 0.5 mL/min was initially selected. The mobile phase comprised a mixture of water (buffer A) and acetonitrile (buffer B) and the gradient utilised is illustrated in **Table 12**.

Table 12 Gradient utilised for the separation of compound **29** and the product

Gradient step	Time (min)	Buffer A (%)	Buffer B (%)
1	1	90	10
2	12	90 to 40	10 to 60
3	14	40 to 95	60 to 5
4	16	95	5

The elution system was utilised each day to monitor the process of the reaction equilibrium. The results are illustrated **Figure 62**.

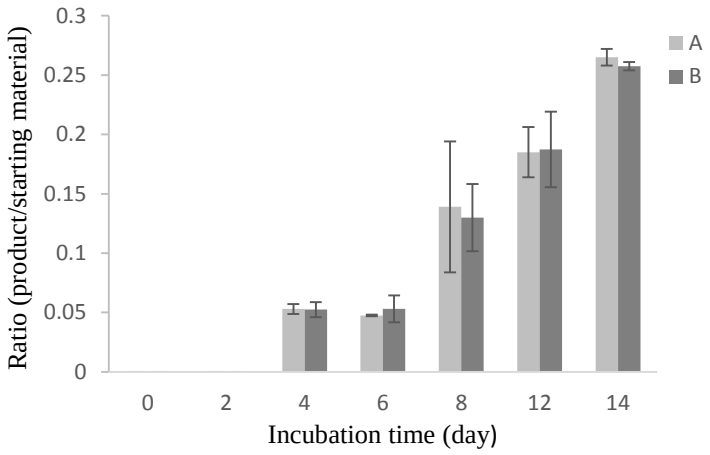


Figure 62 DCC experiment results of compound **29** and **H-1** in the presence or absence of enzyme. Series A: reversible reaction incubated in the absence of enzyme; Series B: reversible reaction incubated in the presence of enzyme. By comparing the ratio of product/starting material, the formation of hydrazone product was not amplified on the addition of enzyme

An analysis of the result suggested that the aldehyde **29** had the peak with retention time (t_R) of around 6.3 min and the t_R of the reductive product was around 8.5 min. (Chromatograms are shown in Appendix section Figure **A2**) The formation of product could not be detected until the fourth day. It was speculated that the structure of the tertiary amine of **H-1** might show steric hindrance and slowed down the reaction rate. In the process of reaction, some noisy peaks appeared during the range of t_R from 1 to 3 min. Also, the baseline of the chromatograms was not flat. These noisy signals interfered with the analysis of the distribution of the DCL products. Some white precipitation was observed in the bottom of Eppendorf in the process of incubation and this was thought to be the inactive enzyme precipitated from the buffer. The reaction solution was used with the HPLC without being pre-treated in previous experiments. Thus, the inactive enzyme precipitation may also have been injected into the analytical instrument and led to the messy peak signal. Thus, the

reaction solution was pre-treated before being applied to the HPLC in order to avoid these undesirable peaks. Samples were removed from the Eppendorf, followed by the addition of methanol and these mixtures were centrifuged for 15 min at 1000 rpm. The supernatants were then used for an HPLC analysis, facilitating the removal of impurities, like the inactive enzyme. This approach was attempted for the HPLC analysis by utilizing the reaction solution containing **29**, **H-1** and enzyme. The noisy peaks from impurity were also greatly reduced. (**Figure 63**)

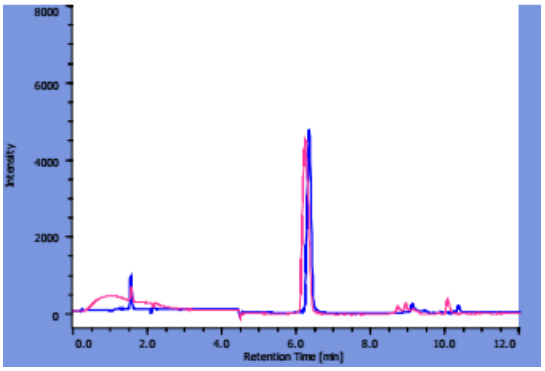


Figure 63 Chromatograms of sample of reaction solution (including **29**, **H-1** and enzyme). Samples were removed from the solution in Eppendorf, followed by the addition of methanol. The mixtures were centrifuged for 15 min at 1000 rpm. The supernatants (100 μL) were then used for the HPLC analysis. The impurity signal appeared between t_R 1 to 3 min reduced greatly.

The chromatograms showed that the compound **29** peak decreased after incubation for 14 days and an extra broad peak appeared with t_R around 4.0 min. Since an excess of the reductive reagent, $NaBH_3CN$, was used to enable the reductive step, it was hypothesised that **29** may be reduced to produce 5-indole uridine and the peak of reductive product 5-indole uridine may appear with t_R (4.0 min). In order to test this hypothesis, a 5-indole uridine sample was utilised with the HPLC under the same analytical conditions and a peak with t_R (4.0 min) appeared in the chromatogram, thus supporting the hypothesis. According to the results of the reversible reaction between compound **29** and **H-1**, the peak of the reductive

product, 5-indole uridine, did not appear until the tenth day of incubation, which suggested that the reductive reaction of the aldehyde started from the tenth day.

Having developed a robust HPLC method for analysis DCC experiment of **H-1** and aldehyde **29** and with these optimised conditions in hand, all the other hydrazide/hydrazine candidates (**H-2** to **H-13**, **Figure 58**) were employed in the reversible reaction to identify amplified library members. The distributions of most DCL members showed no change with the addition of the enzyme template. To our delight, the distribution of one library member formed by compound **29** and 4-(2-hydrazinylethyl)morpholine, **H-9**, was observed to be amplified due to the addition of enzyme. When analysing the chromatograms of reactions in the presence or absence of enzyme, the peak of the product formed increased with the addition of enzyme. It was observed that for, this reversible reaction, peak of the aldehyde building block **29** became broad and overlapped with the product peak. It was suggested that the pH of the reaction solution might change in the process of incubation, which might influence the HPLC results. Thus, the HPLC elution system was modified with a flow rate of 0.6 mL/min. The mobile phase comprised a mixture of water (Buffer A) and acetonitrile (Buffer B) and a new gradient of elution system was developed (**Table 13**). An analysis *via* the optimised HPLC condition indicated the peaks of compound **29** and corresponding product had been separated. By using the optimised elution system, the peak of **29** had shifted from t_R 6.4 min to 5.8 min and the peak of t_R 7.5 min belonged to the product. (Chromatograms are included in the Appendix section as **Figure A2**)

Table 13 Gradient utilised for separation of aldehyde **29** and the corresponding product

Gradient step	Time (min)	Buffer A (%)	Buffer B (%)
1	1	90	10
2	15	90 to 30	10 to 70
3	16	30 to 95	70 to 5

The results of the reversible reaction between **29** and **H-9** are shown in **Figure 64**. The product of reversible reaction was formed from the second incubation day. The rate of reversible reaction with **H-9** was faster than that between **29** and **H-1**. It was suggested that the secondary amine of **H-9** had a much less steric hindrance effect. When comparing the ratio of product/starting material, it was shown that the formation of the product was amplified in the presence of enzyme template from the fourth day and the amplification effect increased during the process of incubation. Therefore, it was suggested that this DCL member could exist binding affinity with target enzyme.

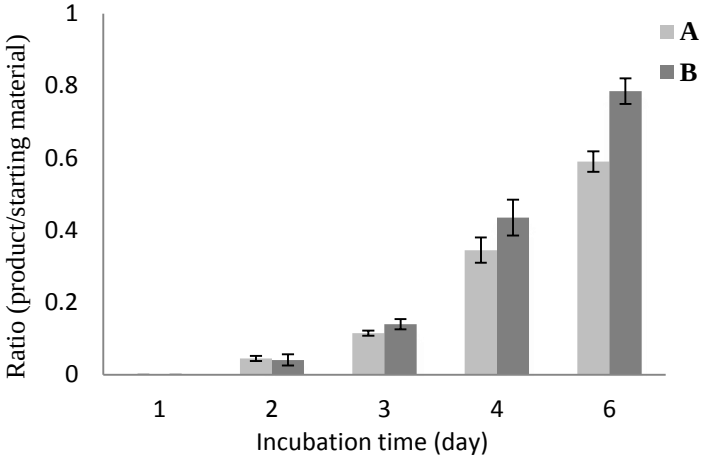
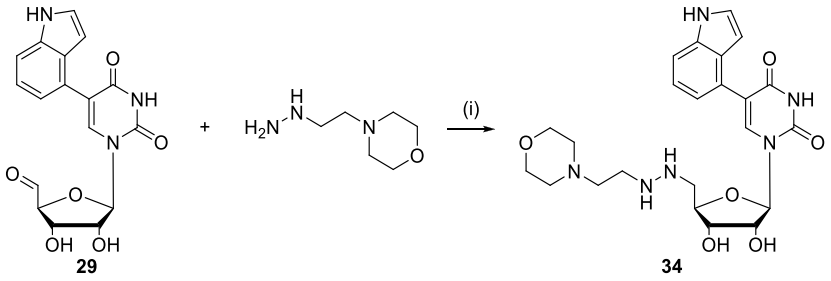


Figure 64 DCC experiment results of compound **29** and **H-9** in the presence or absence of enzyme. Series A: reversible reaction incubated in the absence of enzyme; Series B: reversible reaction incubated in the presence of enzyme. The ratio of product/stating material was calculated based on the corresponding peak areas. When comparing the ratio of product/starting material, the formation of the hydrazone product was amplified with the addition of enzyme

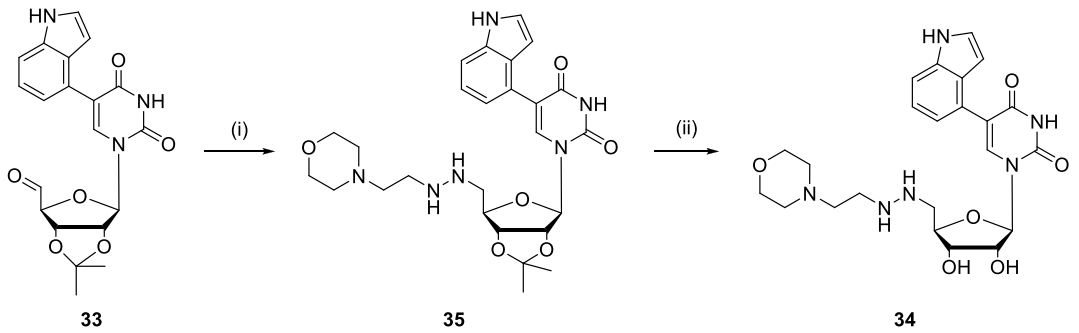
3.3.3. Re-synthesis and activity measurement of active DCL member

Having identified a promising candidate **H-9** from the hydrazide/hydrazine library, the target molecule **34** was synthesised on a larger scale for the measurement of inhibitory activity. (**Scheme 13**) The reaction was carried out using aldehyde **29** and **H-9** in a MeOH solution at room temperature, followed by the addition of $NaBH_3CN$. The desired product was obtained after twice reverse phase chromatography, with a yield of 14%.



Scheme 13 Synthesis of 5-indole uridine hydrazine **34**. Reagents and conditions: (i) $NaBH_3CN$, MeOH, rt, 24h, 14%

Another strategy was employed due to the low yield, in which the coupling reaction with hydrazine **H-9** was carried out prior to the deprotection step. It was hypothesised that the chromatography purification of **35** would be easier to carry out due to its better hydrophobicity, leading to less loss of material during the chromatography. No column purification was required in the deprotection step either. In order to test this hypothesis, the synthesis was initiated from compound **33**. (**Scheme 14**) The coupling reaction was carried out by utilising **33** as the starting material, with the desired product **35** obtained after the reverse phase chromatography purification. Then deprotection was performed in 50% TFA solution at room temperature for 1h. The TLC results suggested that the compound **35** had been decomposed. Therefore, a TFA solution with a lower concentration (20%) was utilised in order to avoid the loss of this starting material, and product **34** was obtained without further purification after removing excess acid solution.



Scheme 14 Synthesis of 5-indole uridine hydrazine. Reagents and conditions: (i) hydrazine **H-9**, $NaBH_3CN$, MeOH, rt, 24h. (ii) 20% TFA, 3h, rt. 25 % in two steps.

With this product in hand, its inhibitory activity was assessed in the biochemical phosphatase-coupled glycosyltransferase assay. While the 5-indole nucleoside **13** inhibited β-1,4-GalTs with an IC_{50} value of around 207 μM, the corresponding derivative **34**, 5-indole nucleoside with hydrazine as a mimic of pyrophosphate and galactose moieties, was inactive at concentrations up to 1 mM. (**Figure 65**)

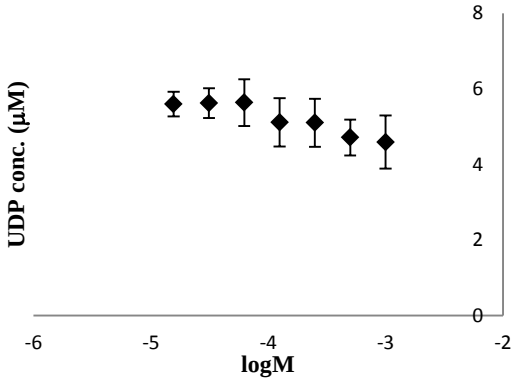


Figure 65 Compound **34** was assessed in the phosphatase-coupled glycosyltransferase assay. No inhibitory activity was observed against β-1,4-GalTs at concentration up to 1 mM.

The distribution of library member **36** (**Figure 66**) was amplified in the presence of β-1,4-GalTs, which suggested that the amplification effects might have resulted from the specific interaction between the enzyme and **36**. Compound **34** was the final stabilized library member which was reduced from **36**, being converted from a hydrogen acceptor into a

hydrogen donor. It might not have maintained the enzyme binding properties of the hydrazone components. It was hypothesised that the hydrazine moiety in the 5'-position of nucleoside was highly flexible due to the reduction of the C=N bond and the hydrazine substituent might have provided steric hindrance which interfered with the enzyme attachment. This interference might have deactivated the inhibition of **34** towards enzyme. The synthesis of non-reduced compound **36** was carried out in order to testify this hypothesis. The related amide analogue **37** was also designed as it contained the C=O linkage, which might be less flexible than the C-N bond of **34**, thereby facilitating the interaction towards enzyme. (**Figure 66**)

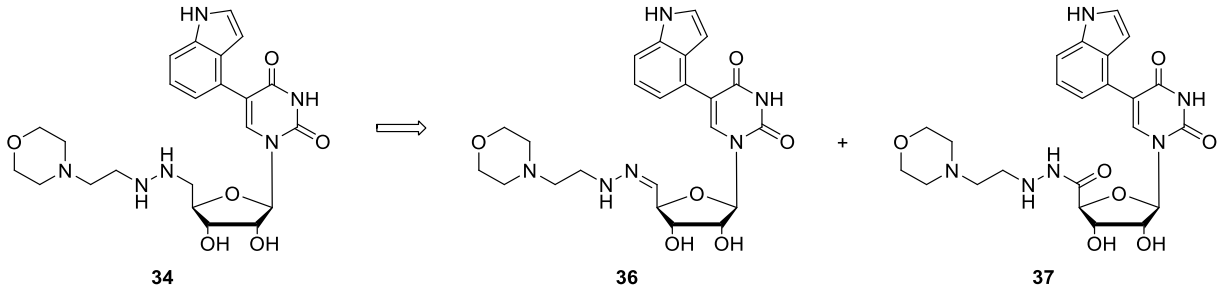
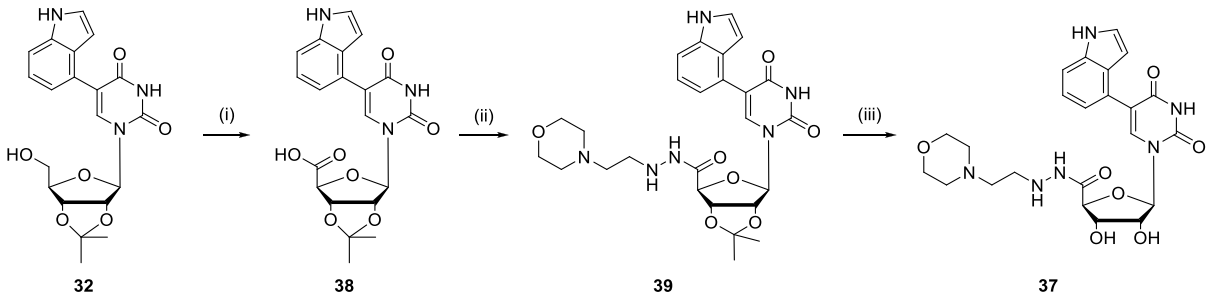


Figure 66 Design of hydrazone **36** and its amide analogue **37**

The attempt of the coupling reaction for the synthesis of compound **36** was made under the same experimental conditions as compound **34** except for the addition of $NaBH_3CN$. After being stirred at room temperature for 24 h, the TLC monitoring indicated that new spots had appeared and there was some unreacted starting material left. Therefore, the mixture was applied to column chromatography for purification. However, no product was obtained after the purification. It was thought that this compound might be unstable in the stationary phase of column and was decomposed in the process of purification. Different chromatography conditions were utilised for **36** purification. For example, reverse phase chromatography with

short column was performed to reduce the purification time; Normal phase chromatographies were also performed with DCM/MeOH and Hexane/EA eluent system separately. However, these all failed and no product **36** was obtained. Thus, the synthesis work was switched to compound **37**.

The synthesis of amide **37** began by utilising protected 5-indole uridine **32** as the starting material. (**Scheme 15**) The condition used to carry out the oxidation involved (2,2,6,6-Tetramethylpiperidine-1-yl)oxyl (TEMPO) (10 mol%), (Diacetoxyiodo)benzene (BAIB) (3 equiv.) in MeCN/water co-solvents. Under this condition, the formation of compound **38** was achieved with a yield of 72%. Then, compound **38** was coupled with hydrazine **H-9,** but the use of HBTU, DIEA at room temperature and DMF as solvent did not result in a good yield. (**Table 14**, **entry 1**) Similarly, compound **38** was coupled with the **H-9** in the presence of EDCI and DMAP in DMF solution, whereas the yield was still very low. (**entry 2**) The attempt to this reaction in the presence of HATU and DIEA in DMF produced a good result, with a satisfying yield of 67%. (**entry 3**) The amide analogue **37** was obtained after deprotection in TFA solution.



Scheme 15 Synthesis of compound **37**. Reagents and conditions: (i) TEMPO, BAIB, MeCN/Water = 1/1, rt, 6h, 72%. (ii) HATU, DIEA, DMF, rt, 1d, 67 %. (iii) 20% TFA, 3h, rt. 47%.

Table 14 Attempts for synthesis of amide **39**

Entry	Reagents	Temp/Time	Results
1	HBTU, DIEA, DMF	Room temperature, 1d	< 10 %
2	EDCI, DMAP, DMF	Room temperature, o/n	Low yield
3	HATU, DIEA, DMF	Room temperature, 1d	67%

With compound **37** in hand, its inhibitory activity was assessed in the phosphatase-coupled glycosyltransferase assay, and the result is shown in **Figure 67**. The assay was carried out as previously used and the maximum concentration was up to 1 mM. In contrast to **34**, the amide **37** showed promising activity towards β-1,4-GalTs, with an IC_{50} value of 178 μM. Compared with compound **13** (5-indole uridine) which showed inhibitory activity around 207 μM, there was a relative slightly improvement in the inhibitory activity of amide **37** and the IC_{50} value was still in the same order.

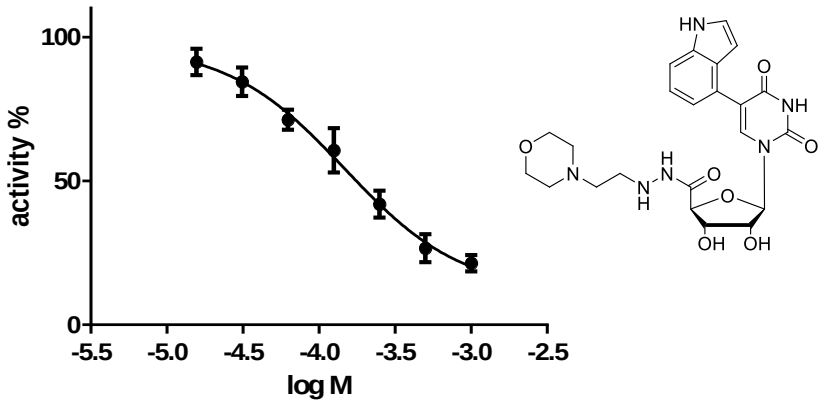
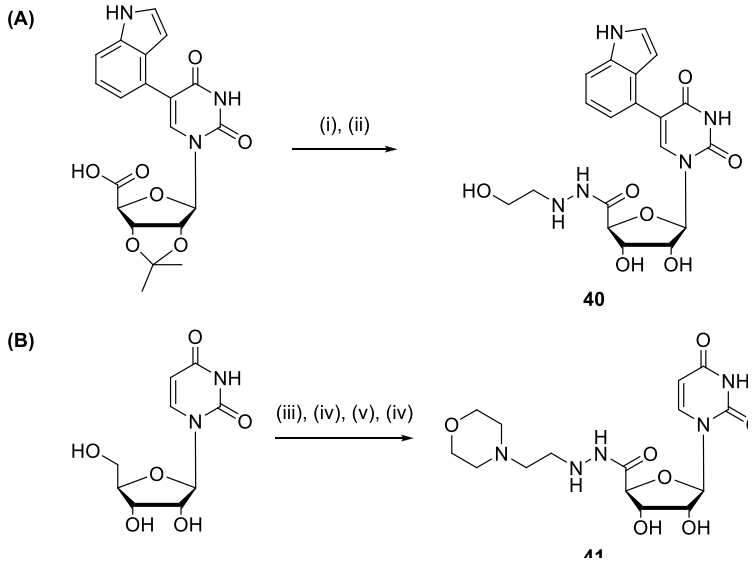


Figure 67 IC_{50} assessment of amide **37** against β-1,4-GalTs. Conditions: **37** (0 to 1 mM) β-1,4-GalTs (diluted to the required concentrations), acceptor (GlcNAc, 5 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (CEL, 1 mg/mL), calf-intestinal alkaline phosphate (CIP, 10 U/mL), DMSO (10%) buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 ºC with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments

Since the activity of amide **37** was similar to that of the nucleotide **13**, it was speculated that the main origin of the binding affinity between amide **37** and enzyme was still the nucleoside

moiety, especially from the indole motif. The addition of hydrazine, which was applied as the mimic of pyrophosphate and galactose, did not give extra binding affinity. One of the inactive hydrazine, **H-11**, was selected and utilised for the synthesis of compound **40** to test the contribution of the hydrazine substituent in the 5'-position of nucleoside to enzyme inhibition. We also prepare the amide derivative **41** without indole group in order to understand the contribution of the indole moiety of compound **37** to enzyme inhibition. (**Scheme 16**)



Scheme 16 Synthesis of amide **40** and **41**. Reagents and conditions: (i) HATU, DIEA, DMF, rt, 1d, 58%. (ii) 20 % TFA, 3h, rt, 43%. (iii) 2,2-dimethoxypropane, *p*-TSA, THF, 60 °C, 3h, 70%. (iv) TEMPO, BIAB, MeCN/Water = 1/1, rt, 6h, 88%. (v) **H-11**, HATU, DIEA, DMF, rt, 1d, 75%. (vi) 20% TFA, 3h, rt. 83%.

As for the synthesis of compound **40**, the acid **38** was coupled with the aliphatic hydrazine **H-11** and the product **40** was obtained after deprotection of diol. (**Scheme 16**, **A**) The synthesis of amide **41** began from the DMP protection of uridine and the DMP-protected uridine was obtained in a yield of 70% after purification. The following step was performed by subjecting the DMP-protected uridine to the condition of TEMPO oxidation, followed by a coupling reaction with hydrazine **H-9**. Then, the deprotected reaction was carried out in a 20 % TFA solution, thereby affording the desired amide product **41**. (**Scheme 16**, **B**)

With these two products in hand, their inhibitory activities were assessed in the phosphatase-coupled glycosyltransferase assay, and the results are shown in **Figure 68**. Compound **41** was inactive towards the β-1,4-GalTs without the indole substituent. In contrast to compound **41**, amide **40** showed an activity towards β-1,4-GalTs, with an IC_{50} value of 213 μM.

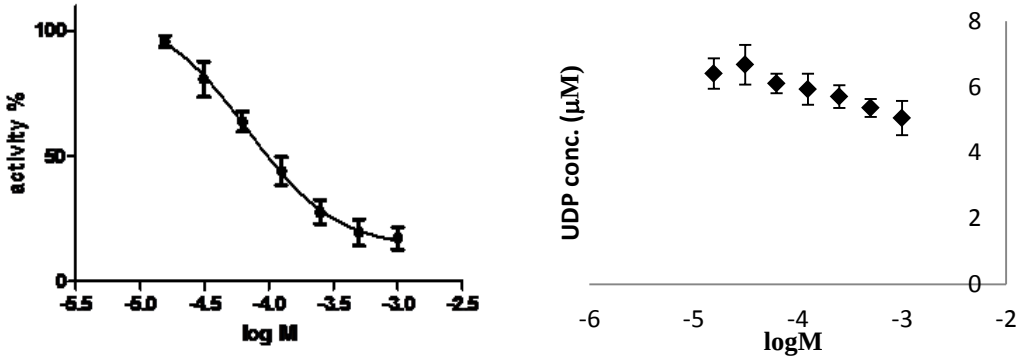
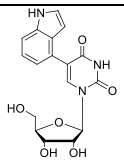
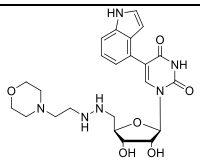
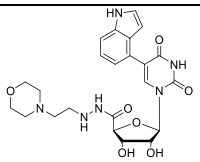
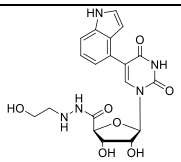
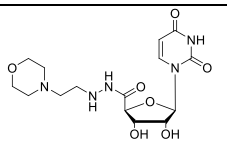


Figure 68 Measurement of inhibitory activities of compound **40** (graph on left side) and **41** (graph on the right side) against β-1,4-GalTs. Conditions: **40** or **41** (0 to 1 mM) β-1,4-GalTs (diluted to the required concentrations), acceptor (GlcNAc, 5 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (CEL, 1 mg/mL), calf-intestinal alkaline phosphate (CIP, 10 U/mL), DMSO (10%) buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 °C with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments

According to previous result (chapter 2), it was suggested that the nucleoside derivative **13** showed inhibitory activity, with IC_{50} of 207 μM. When the hydrazine **H-9** was introduced in 5'-position of nucleoside with a C-N linkage, the corresponding compound **34** was seen to be inactive towards β-1,4-GalT. When a C=O linkage was utilised as the replacement of the C-N linkage, the corresponding compound **37** was active against the enzyme. Compared with **13**, there was a slightly improvement of inhibitory activity of compound **37** and their IC_{50} values were still of the same order of magnitude. The inhibitory activity of amide **41** completely disappeared due to the removal of indole moiety in 5-position of nucleoside. When the hydrazine **H-9**, which was selected as the amplified building block from DCC experiment,

was replaced by an inactive hydrazine **H-11**, the resulting product **40** likewise exhibited inhibitory activity towards β-1,4-GalTs, with a similar IC_{50} value of 213 μM. (**Table 15**)

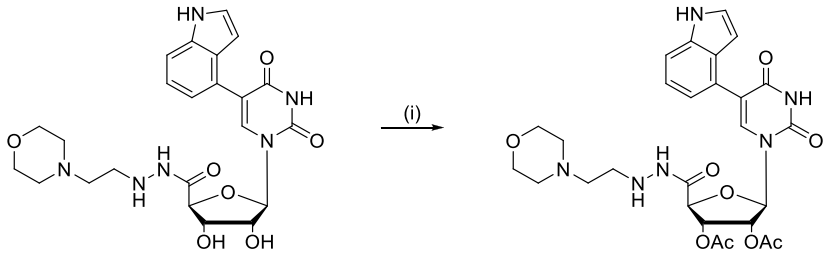
Table 15 Summary of inhibitory activities of nucleoside derivative, DCL members as well as the amide analogue (error means values ± S.D)

Compound	**13**	**34**	**37**	**40**	**41**
Structure					
IC_{50} (μM)	207 ± 37	NO inhibition	178 ± 21	213 ± 16	NO inhibition

The nucleoside building block was responsible for directing compounds towards enzyme active site and the 5-position indole moiety afforded dominant effect for binding affinity towards enzyme. When the hydrazine was introduced to 5'-position of nucleoside *via* C-N linkage as the mimic of pyrophosphate and galactose moieties, the hydrazine moiety could exhibit high flexibility due to the single bond of C-N. This hydrazine tail might provide steric hindrance, which interfered with the enzyme attachment of nucleoside moiety. This interference deactivated the inhibition of **34** towards enzyme. When the C=O was utilised as the linkage, the flexibility of hydrazine substituent in 5'-position reduced and the nucleoside moiety could bind towards enzyme again. Thus, the corresponding amide product **37** could inhibit β-1,4-GalTs. Compared the IC_{50} value of compound **13**, there was no obviously change of IC_{50} values of compound **37** and **40**. It was speculated that the indole moiety of nucleoside still contributed dominant effect for enzyme inhibition while the hydrazines showed very weak binding affinity towards enzyme. This hypothesis was supported by the result that compound **41** was inactive due to the removal of indole moiety.

Cell assay of amide analogue 3.14

As amide **37** exhibited inhibitory activity towards β-1,4-GalTs in the enzyme assay, its pro-drug **42** (**Scheme 17**) was utilized in the cell assay.



Scheme 17 Synthesis of prodrug **42**. Reagents and conditions: Ac_2O, pyridine, DMAP, rt, 6h, 69%.

The MTT cell viability assay was carried out and the data showed that cell viability was not affected by compound **42** up to concentration of 100 μM. (**Figure 69**) The effect of the inhibitor on PSGL-1 levels has not yet been tested because of the time constraints. This assay would be the most important experiment in the next step.

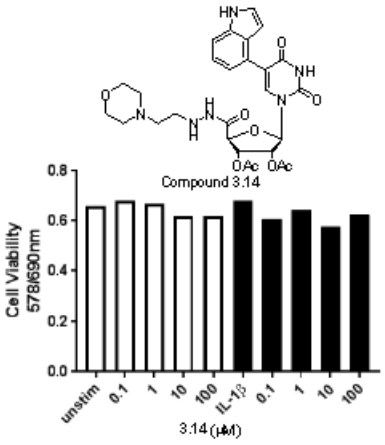


Figure 69 The data of MTT cell viability assay of compound **42**.

3.4. Summary and conclusion.

In conclusion, a hydrazone dynamic combinatorial library (DCL) was designed *via* nucleoside analogue **29** and a collection of hydrazines/hydrazides to identify replacement of pyrophosphate and sugar moieties. The library member generated by hydrazine **H-9** and **29** was identified as the β-1,4-GalTs binder from the library and the corresponding nucleoside derivatives, both reductive product **34** and amide analogue **37**, were synthesized and evaluated in the phosphatase-coupled assay. Only amide **37** was active towards β-1,4-GalTs. It was hypothesized that the linkage between hydrazine substituent and nucleoside could affect the attachment towards enzyme. Compared with the hydrazine substituent, the 5-substituted indole motif of nucleoside contributed dominant effect for enzyme inhibition.

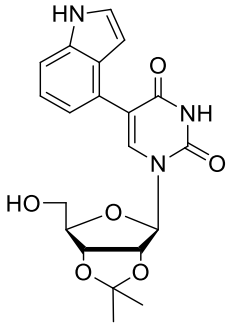
3.5. Experimental section.

General. All chemical reagents were obtained commercially and used as received, unless stated otherwise. Thin layer chromatography (TLC) was performed on pre-coated plates of Silica Gel 60 F254 (Merck), with DCM/MeOH as the mobile phase, unless otherwise stated. Spots were visualised under UV light (254/365nm). NMR spectra were recorded at 300 K on a Bruker BioSpin machine at, respectively, 400.13 MHz (^{1}H-NMR) and 100.62 MHz (^{13}C-NMR). Chemicals shifts (δ) are reported in ppm (parts per million) and coupling constants (J)

in Hz. Mass spectra were recorded at the EPSRC National Mass Spectrometry Facility in Swansea. All yields (%) are isolated yields.

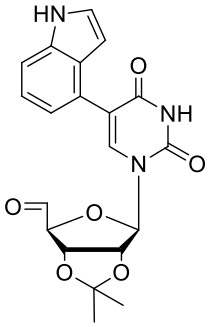
Column chromatography. Method A: Preparative reverse-phase chromatography was performed on a Biologic LP chromatography system equipped with a peristaltic pump and a 254 nm UV Optics Module under the following conditions: Ion-pair chromatography was performed using Lichroprep RP-18 resin equilibrated with 0.05 M TEAB (triethylammonium bicarbonate, pH 7.3). Gradient: 0 – 30% MeOH against 0.05 M TEAB over a total volume of 400 mL. Flow rate: 2 mL/min. Product-containing fractions were combined and repeatedly co-evaporated with methanol to remove residual TEAB.

Method B: Preparative chromatography was carried out on silica gel 60 (pore size 60 Å, 230-400 mesh, Sigma-Aldrich) at normal pressure. DCM and methanol were utilized as mobile phase. Product-containing fractions were combined and evaporated to afford pure products.



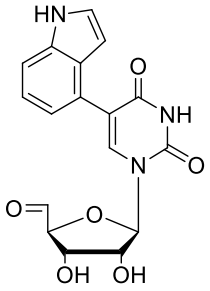
1-((3aR,4R,6R,6aR)-6-(hydroxymethyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-5-(1H-indol-4-yl)pyrimidine-2,4(1H,3H)-dione (32). To a solution of 5-indole uridine (1 equiv., 200 mg) in THF was added 2,2-dimethoxypropane (3 equiv., 1.2 mL) and *p*-TSA (0.05 equiv., 13.2 mg). The reaction mixture was stirred at 60 ^{o}C for 3h. Upon completion of

the reaction, the organic solvent was concentrated and the residue was purified by column chromatography (DCM/MeOH = 30/1), to yield 148 mg of a white form (69%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ7.67 (s, 1H), 7.33 (d, *J* = 12.8 Hz), 7.26 (dd, *J* = 5.6, 12.8 Hz, 1H) 7.15 (dd, *J* = 0.8, 7.2 Hz, 1H), 7.10 (d, *J* = 3.2 Hz, 1H), 6.41 (dd, *J* = 0.8, 2.8 Hz, 1H0, 5.69 (d, *J* = 3.2 Hz, 1H), 5.12 (dd, *J* = 3.2, 6.4 Hz, 1H), 4.98 (dd, *J* = 3.2, 6.4 Hz, 1H), 4.31 (m, 1H), 3.90 (m, 1H), 3.78 (m, 1H), 1.59 (s, 3H), 1.38 (s, 3H). ^{13}C-NMR (100 MHz, $CDCl_3$): δ168.0 (C-4), 155.0 (C-2), 135.6, 131.2, 126.9, 125.9, 124.0, 123.3, 122.3, 112.3, 116.8 (C-6 + C-indole), 113.0 (C-5), 102.4 (C-indole-3), 102.0 (C-1'), 88.2 (C-4'), 84.9 (C-2'), 82.3 (C-3'), 63.9 (C-5'), 31.2 (C-Me), 31.0 (C-Me). m/z (ESI) 398.1352 $[M-H]^-$, $C_{20}H_{20}N_3O_6$ require 398.1357.

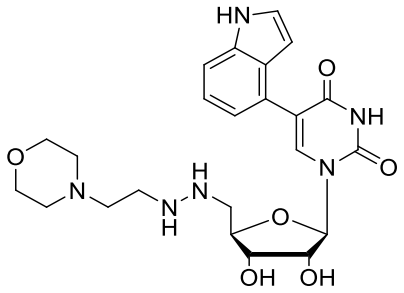


(3aS,4S,6R,6aR)-6-(5-(1H-indol-4-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxole-4-carbaldehyde (33). To a solution of protected 5-indole uridine (1 equiv., 50 mg) in DCM was added dess-martin reagent (3 equiv., 167 mg). The reaction mixture was stirred at rt for 5h and the progress of the reaction was monitored by TLC (DCM/MeOH = 10/1). Upon completion of the reaction, the insoluble powder was removed after filtration, which was followed by the concentration. The residue was purified by column chromatography (DCM/MeOH = 40/1), to yield 22.8 mg of a white form (47%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ9.37 (s, 1H), 7.93 (s, 1H), 7.35 (d, *J* = 8 Hz),

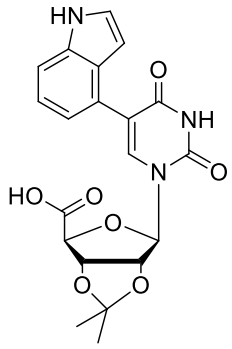
7.24 (d, *J* = 3.2 Hz), 7.13 (t, *J* = 8 Hz, 1H), 7.08 (dd, *J* = 1.2, 7.2 Hz), 6.42 (dd, *J* = 0.8, 3.2 Hz, 1s), 6.04 (d, *J* = 5.2 Hz, 1s), 5.36 (dd, *J* = 3.2, 6.4 Hz, 1H), 5.20 (dd, *J* = 3.2, 6.4 Hz, 1H). 4.96 (m, 1H). 1.50 (s, 3H), 1.39 (s, 3H). ^{13}C-NMR (100 MHz, $CDCl_3$): δ190.8 (C-CHO), 166.9 (C-4), 155.8 (C-2), 139.0, 134.2, 127.6, 124.6, 124.0, 123.2, 122.9, 113.6, 112.6 (C-6 + C-indole), 110.9 (C-5), 103.4 (C-indole-3), 102.7 (C-1'), 87.3 (C-4'), 85.4 (C-2'), 76.5 (C-3'), 31.0 (C-Me), 31.2 (C-Me). m/z (ESI) 396.1196 $[M-H]^-$ $C_{20}H_{18}N_3O_6$ requires 396.1207.



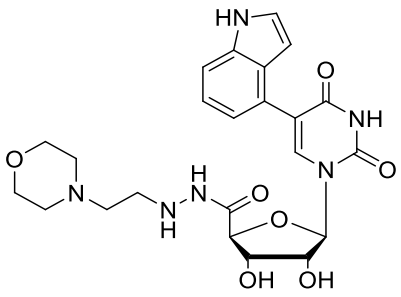
(2S,3S,4R,5R)-5-(5-(1H-indol-4-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-3,4-dihydroxytetrahydrofuran-2-carbaldehyde (29). Protected 5-indole uridine aldehyde (1 equiv, 80 mg) was dissolved in 50 % TFA solution and the mixture was stirred at rt for 30 min and the progress was monitored by TLC (DCM/MeOH = 10/1). The product was obtained after evaporation without further purification, to yield 60 mg of a white form (80%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ9.56 (s, 1H), 7.97 (s, 1H), 7.45 (d, *J* = 8 Hz), 7.37 (d, *J* = 3.2 Hz), 7.13 (t, *J* = 8 Hz, 1H), 7.10 (dd, *J* = 1.2, 7.6 Hz), 6.44 (dd, *J* = 0.8, 3.6 Hz, 1s), 5.99 (d, *J* = 5.2 Hz, 1s), 5.39 (dd, *J* = 3.2, 6.4 Hz, 1H), 5.27 (dd, *J* = 3.2, 6.4 Hz, 1H), 4.67 (m, 1H). ^{13}C-NMR (100 MHz, MeOD): δ 200.0 (C-COOH), 169.3 (C-4), 153.0 (C-2), 139.2, 132.9, 129.4, 124.6, 127.6, 123.9, 119.7, 115.0 (C-6 + C-indole), 110.0 (C-5), 103.4 (C-indole-3), 102.0 (C-1), 87.3 (C-4'), 87.4 (C-2'), 85.6 (C-3'). m/z (ESI) 356.0883 $[M-H]^-$ $C_{17}H_{14}N_3O_6$ requires 356.0889.



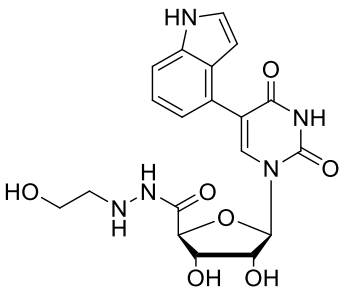
1-((2R,3R,4S,5R)-3,4-dihydroxy-5-((2-(2-morpholinoethyl)hydrazinyl)methyl)tetrahydrofuran-2-yl)-5-(1H-indol-4-yl)pyrimidine-2,4(1H,3H)-dione (34). To the solution of 5-indole uridine aldehyde (1 equiv., 22 mg) in MeOH was added 4-(2-hydrazinylethyl) morpholine (1 equiv., 6.7 mg). The reaction mixture was stirred at rt for 1d, which was followed by addition of $NaBH_3CN$ (10 mol%). Upon of completion of the reaction, the residue was concentrated and purified by reverse phase chromatography. The obtained pure product was dissolved in 20 % TFA solution at rt for 3h and the progress was monitored by TLC (DCM/MeOH = 4/1) The product was obtained after evaporation without further purification, to yield 4.7 mg of light yellow form (26%) in two steps. ^{1}H-NMR (400 MHz, MeOD) δ 7.93 (s, 1H), 7.45 (d, *J* = 8 Hz, 1H), 7.35 (d, *J* = 3.6 Hz), 7.20 (t, *J* = 7.2 Hz, 1H), 7.16 (dd, *J* = 1.2, 7.2 Hz, 1H), 6.52 (dd, *J* = 0.8, 7.2 Hz, 1H), 6.00 (d, *J* = 5.2 Hz, 1H), 4.51 (dd, *J* = 5.2, 8.8 Hz, 1H), 4.30 (dd, *J* = 5.2, 8.4 Hz, 1H), 4.06 (m, 1H), 3.82 (m, 2H), 3.70 (m, 2H), 3.52 (m, 2H), 3.24 (m, 2H), 3.21 (dd, *J* = 2.8, 7.2 Hz, 1H), 3.15 (dd, *J* = 3.2, 7.2 Hz, 1H). ^{13}C-NMR (100 MHz, MeOD): δ 170.0 (C-4), 155.5 (C-2), 146.3, 137.2, 133.0, 125.4, 124.3, 123.7, 116.7, 115.8 (C-6 + C-indole), 114.0 (C-5), 102.9 (C-indole-3), 100.2 (C-1'), 89.6 (C-4'), 87.7 (C-2'), 86.0 (C- 2'), 79.0 (C-3'), 68.5 (C-morpholine), 68.0 (C-morpholine), 59.4 (C-morpholine), 58.8 (C-morpholine), 56.3 (C-5'), 52.0 (C-4''), 50.0 (C-3''). m/z (ESI) 485.2149 $[M-H]^-$ $C_{23}H_{29}N_6O_6$ requires 485.2133.



(3aS,4S,6R,6aR)-6-(5-(1H-indol-4-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxole-4-carboxylic acid (38). To a solution of protected 5-indole uridine (1 equiv., 220 mg) in MeCN/Water = 1/1 was added TEMPO (10 mol%, 24 mg), BIAB (2.2 equiv., 548 mg). The reaction mixture was stirred at rt for 6h and TLC was used to monitor the progress. Upon of the completion of the reaction, the reaction solution was diluted with Et_2O, which was followed by filtration, to yield 144 mg product of light brown powder (65%). ^{1}H-NMR (400 MHz, MeOD) δ7.77 (s, 1H), 7.40 (d, J = 8 Hz, 1H), 7.39 (d, J = 4 Hz, 1H), 7.29 (t, J = 7.6 Hz, 1H), 7.12 (dd, J = 1.2, 6.8 Hz, 1H), 6.43 (dd, J = 1.2, 3.2 Hz, 1H), 5.71 (s, 1H), 5.37 (dd, J = 2, 6 Hz, 1H), 5.26 (dd, J = 2, 6.4 Hz, 1H), 4.67 (d, J = 1.6 Hz, 1H), 1.56 (s, 3H), 1.40 (s, 3H). ^{13}C-NMR (100 MHz, MeOD): δ173.6 (C-COOH), 166.5 (C-4), 152.6 (C-2), 146.8, 139.9, 134.9, 127.8, 127.2, 124.3, 114.3, 113.3 (C-6 + C-indole), 111.1 (C-5), 102.5 (C-indole-3), 102.2 (C-1’), 89.2 (C-4’), 86.1 (C-2’), 85.7 (C-3’), 27.0 (C-Me), 25.2 (C-Me). m/z (ESI) 412.1145 $[M-H]^-$ $C_{20}H_{18}N_3O_7$ requires 412.1140.



(2S,3S,4R,5R)-5-(5-(1H-indol-4-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-3,4-dihydroxy-N'-(2-morpholinoethyl)tetrahydrofuran-2-carbohydrazide (37). To a solution of protected 5-indole uridine (1 equiv., 50 mg) acid in DMF, 4-(2-hydrazinylethyl)morpholine (1 equiv., 30 mg), HATU (2 equiv., 108 mg), DIEA (3 equiv., 155 mg) was added. The reaction mixture was stirred at rt for 24 h and TLC was used to monitor the progress of reaction. Upon of completion of the reaction, the residue was purified by reverse phase chromatography. The obtained pure product was dissolved in 20% TFA solution and the mixture was stirred at rt for 3h. The de-protected product was obtained after concentration without further purification, to yield 16 mg of 24% in two steps. ^{1}H-NMR (400 MHz, MeOD) δ 7.87 (s, 1H), 7.46 (d, J = 8 Hz, 1H), 7.35 (d, J = 3.6 Hz, 1H), 7.24 (t, J = 7.2 Hz, 1H), 7.09 (dd, J = 1.2, 6.8 Hz, 1H), 6.43 (dd, J = 1.2, 3.2 Hz, 1H), 6.03 (d, J = 2 Hz, 1H), 5.29 (m, 1H), 5.12 (dd, J = 1.6, 6 Hz), 4.97 (dd, J = 2, 6 Hz, 1H), 3.82 (m, 2H), 3.70 (m, 2H), 3.52 (m, 2H), 3.24 (m, 2H). ^{13}C-NMR (100 MHz, MeOD): δ 179.0 (C-CONH), 163.9 (C-4), 159.8 (C-2), 149.9, 144.9, 138.6, 129.0, 127.8, 126.4, 117.1, 113.2 (C-6 + C-indole), 112.7 (C-5), 105.0 (C-indole-3), 99.6 (C-1'), 83.1 (C-4'), 82.4 (C-2'), 75.8 (C-3'), 69.9 (C-morpholine), 68.7 (C-morpholine), 60.8 (C-morpholine), 60.0 (C-morpholine), 59.3 (C-4"), 53.2 (C-3"). m/z (ESI) 499.1941 $[M-H]^{-}$ $C_{23}H_{27}N_6O_7$ requires 499.1936.



(2S,3S,4R,5R)-5-(5-(1H-indol-4-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-3,4-dihydroxy-N'-(2-hydroxyethyl)tetrahydrofuran-2-carbohydrazide (41). To a solution of

protected 5-indole uridine (1 equiv., 50 mg) acid in DMF, 2-hydrazinylethanol (1 equiv., 11 mg), HATU (2 equiv., 108 mg), DIEA (3 equiv., 155 mg) was added. The reaction mixture was stirred at rt for 24 h and TLC was used to monitor the progress of reaction. Upon of completion of the reaction, the residue was purified by reverse phase chromatography. The obtained pure product was dissolved in 20% TFA solution and the mixture was stirred at rt for 3h. The de-protected product was obtained after concentration without further purification, to yield 12 mg of 25% in two steps. ^{1}H-NMR (400 MHz, MeOD) δ7.87 (s, 1H), 7.46 (d, *J* = 8 Hz, 1H), 7.35 (d, *J* = 3.6 Hz, 1H), 7.24 (t, *J* = 7.2 Hz, 1H), 7.09 (dd, *J* = 1.2, 6.8 Hz, 1H), 6.43 (dd, *J* = 1.2, 3.2 Hz, 1H), 5.81 (d, *J* = 2 Hz, 1H), 5.37 (m, 1H), 5.08 (dd, *J* = 1.6, 6 Hz), 4.97 (dd, *J* = 2, 6 Hz, 1H), 3.65 (m, 2H), 3.58 (m, 1H), 3.50 (m, 1H). ^{13}C-NMR (100 MHz, MeOD): δ178.0 (C-CONH), 163.9 (C-4), 158.0 (C-2), 150.0, 147.3, 136.9, 129.6, 126.6, 126.4, 118.6, 116.2 (C-6 + C-indole), 111.9 (C-5), 106.0 (C-indole-3), 100.6 (C-1'), 85.6 (C-4'), 84.6 (C-2'), 74.3 (C-3'), 60.6 (C-4"), 53.2 (C-3"). m/z (ESI) 430.1363 $[M-H]^-$ $C_{19}H_{20}N_5O_7$ requires 430.1375.

GalT Inhibition experiments.

General. All reagents for the biochemical assays were obtained commercially and used as received, unless otherwise stated. 5-FT UDP-Gal was prepared as previously reported and used as a positive control. Bovine β-1, 4-GalT was expressed, purified and refolded using an adaptation of a previously reported protocol. LgtC was expressed and purified as previously reported. Inhibition assays were carried out as previously reported and briefly described below. Absorbance measurements were carried out on a BMG Labtech POLARstar Optima multiplate reader.

Data collection and analysis protocol. Assays were carried out on 96-well plates. On each microplate, sample, control and background wells were included in triplicate. A calibration curve (0-12.5 μM UDP, corresponding to 0-25 μM Pi) was constructed for each microplate by linear regression. The calibration curve was used to convert absorbance measurements at 620 nm in sample and control wells to [UDP] (μM). For each sample and control well, a corresponding background well (containing identical components but no acceptor) was included, to account for non-specific hydrolysis of donor. Corrected absorbance values for each well were obtained by subtracting the corresponding background reading form the absorbance of the respective smaple or control well. Inhibition (%) was calculated by dividing absorbance in the presence of inhibitor by maximum absorbance (negative control, no inhibitor). Percentage inhibition was plotted over log[inhibitor] and analysed with GraphPad Prism 6 software to obtain relative IC50 values. Averages and standard deviations were calculated in Microsoft Excel.

DCC experiment Solutions of hydrazines/hydrazides, indole-uridine aldehyde **29** and sodium cyanoborohydride were introduced into 0.5 mL eppendorf containing a solution of β1,4GalTs in MOPS buffer. (Buffer: 50 mM MOPS, 20 $MnCl_2$, pH 7.4) The final concentration of each reagents: hydrazines or hydrazides (300 μL), **29** (300 μL), $NaBH_3CN$ (5 mM), enzyme (0.69 mg/mL). Total volume of each eppendorf is 300 μL. Then the reaction mixture was incubated in a water bath of 25 ºC and monitored by HPLC of a JASCO MD-2018 with the detection at 254 nm. 100 μL samples solution were removed from the Eppendorf, followed by the addition of methanol. The mixture was centrifuged for 15 min at 1000 rpm. The supernatant was then injected on a reverse phase column (Luna 5u C8(2) 100A, new column 150 ×4.6 mm) and eluted at 0.5 mL/min (for gradient A method) or 0.6

mL/min (for gradient B method). The mobile phase comprised a mixture of water (buffer A) and acetonitrile (buffer B) and the gradient A (**Table 12**) or gradient B (**Table 13)** was utilized for the separation and detection.

HPLC gradient A

Table 12 Gradient utilised for separation of aldehyde **29** and the corresponding products

Gradient step	Time (min)	Buffer A (%)	Buffer B (%)
1	1	90	10
2	12	90 to 40	10 to 60
3	14	40 to 95	60 to 5
4	16	95	5

HPLC gradient B

Table 13 Gradient utilised for the separation aldehyde **29** and the corresponding products

Gradient step	Time (min)	Buffer A (%)	Buffer B (%)
1	1	90	10
2	15	90 to 30	10 to 70
3	16	30 to 95	70 to 5

3.6. References

1. Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J. L.; Sanders, J. K.; Otto, S., Dynamic combinatorial chemistry. *Chem Rev* **2006,** *106*, 3652-711.
2. Mondal, M.; Hirsch, A. K., Dynamic combinatorial chemistry: a tool to facilitate the identification of inhibitors for protein targets. *Chem Soc Rev* **2015,** *44*, 2455-88.
3. Huang, R. J.; Leung, I. K. H., Protein-Directed Dynamic Combinatorial Chemistry: A Guide to Protein Ligand and Inhibitor Discovery. *Molecules* **2016,** *21*.
4. Hioki, H.; Still, W. C., Chemical evolution: A model system that selects and amplifies a receptor for the tripeptide (D)Pro(L)Val(D)Val. *J Org Chem* **1998,** *63*, 904-905.
5. Karan, C.; Miller, B. L., RNA-selective coordination complexes identified via dynamic combinatorial chemistry. *J Am Chem Soc* **2001,** *123*, 7455-7456.
6. Epstein, D. M.; Choudhary, S.; Churchill, M. R.; Keil, K. M.; Eliseev, A. V.; Morrow, J. R., Chloroform-soluble Schiff-base Zn(II) or Cd(II) complexes from a dynamic combinatorial library. *Inorg Chem* **2001,** *40*, 1591-1596.
7. Eliseev, A. V.; Nelen, M. I., Use of molecular recognition to drive chemical evolution, Part 2. Mechanisms of an automated genetic algorithm implementation. *Chem-Eur J* **1998,** *4*, 825-834.
8. Huc, I.; Lehn, J. M., Virtual combinatorial libraries: dynamic generation of molecular and supramolecular diversity by self-assembly. *Proc Natl Acad Sci U S A* **1997,** *94*, 2106-10.
9. Shi, B.; Stevenson, R.; Campopiano, D. J.; Greaney, M. F., Discovery of Glutathione S-Transferase Inhibitors Using Dynamic Combinatorial Chemistry. *J Am Chem Soc* **2006,** *128*, 8459-8467.
10. Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H., Structure-Based Design of Inhibitors of the Aspartic Protease Endothiapepsin by Exploiting Dynamic Combinatorial Chemistry. *Angew Chem Int Edit* **2014,** *53*, 3259-3263.
11. Huang, R.; Leung, I., Protein-Directed Dynamic Combinatorial Chemistry: A Guide to Protein Ligand and Inhibitor Discovery. *Molecules* **2016,** *21*, 910.
12. Huang, R.; Leung, I. K., Protein-Directed Dynamic Combinatorial Chemistry: A Guide to Protein Ligand and Inhibitor Discovery. *Molecules* **2016,** *21*.
13. Ramstrom, O.; Lehn, J. M., In situ generation and screening of a dynamic

combinatorial carbohydrate library against concanavalin A. *Chembiochem* **2000,** *1*, 41-8.

14. Nasr, G.; Petit, E.; Supuran, C. T.; Winum, J. Y.; Barboiu, M., Carbonic anhydrase II-induced selection of inhibitors from a dynamic combinatorial library of Schiff's bases. *Bioorg Med Chem Lett* **2009,** *19*, 6014-6017.

15. Leung, I. K. H.; Brown, T.; Schofield, C. J.; Claridge, T. D. W., An approach to enzyme inhibition employing reversible boronate ester formation. *Medchemcomm* **2011,** *2*, 390-395.

16. Schmidt, M. F.; Rademann, J., Dynamic template-assisted strategies in fragment-based drug discovery. *Trends Biotechnol* **2009,** *27*, 512-521.

17. Mondal, M.; Hirsch, A. K. H., Dynamic combinatorial chemistry: a tool to facilitate the identification of inhibitors for protein targets. *Chem Soc Rev* **2015,** *44*, 2455-2488.

18. Hochgurtel, M.; Biesinger, R.; Kroth, H.; Piecha, D.; Hofmann, M. W.; Krause, S.; Schaaf, O.; Nicolau, C.; Eliseev, A. V., Ketones as building blocks for dynamic combinatorial libraries: Highly active neuraminidase inhibitors generated via selection pressure of the biological target. *J Med Chem* **2003,** *46*, 356-358.

19. Poulsen, S. A., Direct screening of a dynamic combinatorial library using mass spectrometry. *J Am Soc Mass Spectr* **2006,** *17*, 1074-1080.

20. Mondal, M.; Radeva, N.; Koster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H., Structure-Based Design of Inhibitors of the Aspartic Protease Endothiapepsin by Exploiting Dynamic Combinatorial Chemistry. *Angew Chem Int Edit* **2014,** *53*, 3259-3263.

21. Lew, W.; Chen, X.; Kim, C. U., Discovery and development of GS 4104 (oseltamivir): an orally active influenza neuraminidase inhibitor. *Curr Med Chem* **2000,** *7*, 663-72.

22. Valade, A.; Urban, D.; Beau, J. M., Two galactosyltransferases' selection of different binders from the same uridine-based dynamic combinatorial library. *J Comb Chem* **2007,** *9*, 1-4.

23. Valade, A.; Urban, D.; Beau, J. M., Target-assisted selection of galactosyltransferase binders from dynamic combinatorial libraries. An unexpected solution with restricted amounts of the enzyme. *Chembiochem* **2006,** *7*, 1023-1027.

24. Biet, T.; Peters, T., Molecular recognition of UDP-Gal by beta-1,4-galactosyltransferase T1. *Angew Chem Int Edit* **2001,** *40*, 4189-4192.

25. Poulsen, S.-A., Direct Screening of a Dynamic Combinatorial Library Using Mass Spectrometry. *J Am Soc Mass Spectr* **2006,** *17*, 1074-1080.

26. Pastorekova, S.; Parkkila, S.; Pastorek, J.; Supuran, C. T., Carbonic anhydrases: Current state of the art, therapeutic applications and future prospects. *J Enzym Inhib Med Ch* **2004,** *19*, 199-229.

27. Zameo, S.; Vauzeilles, B.; Beau, J. M., Direct composition analysis of a dynamic library of imines in an aqueous medium. *Eur J Org Chem* **2006**, 5441-5444.

28. Sindelar, M.; Lutz, T. A.; Petrera, M.; Wanner, K. T., Focused pseudostatic hydrazone libraries screened by mass spectrometry binding assay: optimizing affinities toward gamma-aminobutyric acid transporter 1. *J Med Chem* **2013,** *56*, 1323-40.

29. Caraballo, R.; Sakulsombat, M.; Ramstrom, O., Towards Dynamic Drug Design: Identification and Optimization of beta-Galactosidase Inhibitors from a Dynamic Hemithioacetal System. *Chembiochem* **2010,** *11*, 1600-1606.

30. Caraballo, R.; Dong, H.; Ribeiro, J. P.; Jimenez-Barbero, J.; Ramstrom, O., Direct STD NMR identification of beta-galactosidase inhibitors from a virtual dynamic hemithioacetal system. *Angew Chem Int Ed Engl* **2010,** *49*, 589-93.

31. Leung, I. K. H.; Demetriades, M.; Hardy, A. P.; Lejeune, C.; Smart, T. J.; Szollossi, A.; Kawamura, A.; Schofield, C. J.; Claridge, T. D. W., Reporter Ligand NMR Screening Method for 2-Oxoglutarate Oxygenase Inhibitors. *J Med Chem* **2013,** *56*, 547-555.

32. Poulsen, S. A.; Bornaghi, L. F., Fragment-based drug discovery of carbonic anhydrase II inhibitors by dynamic combinatorial chemistry utilizing alkene cross metathesis. *Bioorgan Med Chem* **2006,** *14*, 3275-3284.

33. Kern, F. T.; Wanner, K. T., Generation and Screening of Oxime Libraries Addressing the Neuronal GABA Transporter GAT1. *Chemmedchem* **2015,** *10*, 396-410.

34. Nasr, G.; Petit, E.; Supuran, C. T.; Winum, J. Y.; Barboiu, M., Carbonic anhydrase II-induced selection of inhibitors from a dynamic combinatorial library of Schiff's bases. *Bioorg Med Chem Lett* **2009,** *19*, 6014-7.

35. Berl, V.; Huc, I.; Lehn, J. M.; DeCian, A.; Fischer, J., Induced fit selection of a barbiturate receptor from a dynamic structural and conformational/configurational library. *Eur J Org Chem* **1999**, 3089-3094.

36. Bornaghi, L. F.; Wilkinson, B. L.; Kiefel, M. J.; Poulsen, S. A., Synthesis of cyclic oligomers of a modified sugar amino acid utilising dynamic combinatorial chemistry. *Tetrahedron Lett* **2004,** *45*, 9281-9284.

37. Nguyen, R.; Huc, I., Optimizing the reversibility of hydrazone formation for dynamic combinatorial chemistry. *Chem Commun* **2003**, 942-943.

38. Congreve, M. S.; Davis, D. J.; Devine, L.; Granata, C.; O'Reilly, M.; Wyatt, P. G.;

Jhoti, H., Detection of ligands from a dynamic combinatorial library by X-ray crystallography. *Angew Chem Int Ed Engl* **2003,** *42*, 4479-82.

39. Woon, E. C.; Demetriades, M.; Bagg, E. A.; Aik, W.; Krylova, S. M.; Ma, J. H.; Chan, M.; Walport, L. J.; Wegman, D. W.; Dack, K. N.; McDonough, M. A.; Krylov, S. N.; Schofield, C. J., Dynamic combinatorial mass spectrometry leads to inhibitors of a 2-oxoglutarate-dependent nucleic acid demethylase. *J Med Chem* **2012,** *55*, 2173-84.

40. Cancilla, M. T.; He, M. M.; Viswanathan, N.; Simmons, R. L.; Taylor, M.; Fung, A. D.; Cao, K.; Erlanson, D. A., Discovery of an Aurora kinase inhibitor through site-specific dynamic combinatorial chemistry. *Bioorg Med Chem Lett* **2008,** *18*, 3978-3981.

41. Scott, D. E.; Dawes, G. J.; Ando, M.; Abell, C.; Ciulli, A., A Fragment-Based Approach to Probing Adenosine Recognition Sites by Using Dynamic Combinatorial Chemistry. *Chembiochem* **2009,** *10*, 2772-2779.

42. Milanesi, L.; Hunter, C. A.; Sedelnikova, S. E.; Waltho, J. P., Amplification of bifunctional ligands for calmodulin from a dynamic combinatorial library. *Chem-Eur J* **2006,** *12*, 1081-1087.

43. Danieli, B.; Giardini, A.; Lesma, G.; Passarella, D.; Peretto, B.; Sacchetti, A.; Silvani, A.; Pratesi, G.; Zunino, F., Thiocolchicine-podophyllotoxin conjugates: Dynamic libraries based on disulfide exchange reaction. *J Org Chem* **2006,** *71*, 2848-2853.

44. Shi, B. L.; Greaney, M. F., Reversible Michael addition of thiols as a new tool for dynamic combinatorial chemistry (pg 886, 2005). *Chem Commun* **2005**, 2181-2181.

45. Shi, B. L.; Stevenson, R.; Campopiano, D. J.; Greaney, M. F., Discovery of glutathione S-transferase inhibitors using dynamic combinatorial chemistry. *J Am Chem Soc* **2006,** *128*, 8459-8467.

46. Misuraca, M. C.; Moulin, E.; Ruff, Y.; Giuseppone, N., Experimental and theoretical methods for the analyses of dynamic combinatorial libraries. *New J Chem* **2014,** *38*, 3336-3349.

47. Valade, A.; Urban, D.; Beau, J. M., Target-assisted selection of galactosyltransferase binders from dynamic combinatorial libraries. An unexpected solution with restricted amounts of the enzyme. *Chembiochem* **2006,** *7*, 1023-7.

48. Woon, E. C. Y.; Demetriades, M.; Bagg, E. A. L.; Aik, W.; Krylova, S. M.; Ma, J. H. Y.; Chan, M. C.; Walport, L. J.; Wegman, D. W.; Dack, K. N.; McDonough, M. A.; Krylov, S. N.; Schofield, C. J., Dynamic Combinatorial Mass Spectrometry Leads to Inhibitors of a 2-Oxoglutarate-Dependent Nucleic Acid Demethylase. *J Med Chem* **2012,** *55*, 2173-2184.

49. Rose, N. R.; Woon, E. C. Y.; Kingham, G. L.; King, O. N. F.; Mecinovic, J.; Clifton, I.

J.; Ng, S. S.; Talib-Hardy, J.; Oppermann, U.; McDonough, M. A.; Schofield, C. J., Selective Inhibitors of the JMJD2 Histone Demethylases: Combined Nondenaturing Mass Spectrometric Screening and Crystallographic Approaches. *J Med Chem* **2010,** *53*, 1810-1818.

50. Hochgurtel, M.; Kroth, H.; Piecha, D.; Hofmann, M. W.; Nicolau, C.; Krause, S.; Schaaf, O.; Sonnenmoser, G.; Eliseev, A. V., Target-induced formation of neuraminidase inhibitors from in vitro virtual combinatorial libraries. *P Natl Acad Sci USA* **2002,** *99*, 3382-3387.

51. Zameo, S.; Vauzeilles, B.; Beau, J. M., Dynamic combinatorial chemistry: Lysozyme selects an aromatic motif that mimics a carbohydrate residue. *Angew Chem Int Edit* **2005,** *44*, 965-969.

52. Poulsen, S. A.; Bornaghi, L. F., Fragment-based drug discovery of carbonic anhydrase II inhibitors by dynamic combinatorial chemistry utilizing alkene cross metathesis. *Bioorg Med Chem* **2006,** *14*, 3275-84.

53. Ruff, Y.; Garavini, V.; Giuseppone, N., Reversible Native Chemical Ligation: A Facile Access to Dynamic Covalent Peptides. *J Am Chem Soc* **2014,** *136*, 6333-6339.

54. Larson, K. K.; He, M.; Teichert, J. F.; Naganawa, A.; Bode, J. W., Chemical sensing with shapeshifting organic molecules. *Chem Sci* **2012,** *3*, 1825-1828.

55. Dal Molin, M.; Gasparini, G.; Scrimin, P.; Rastrelli, F.; Prins, L. J., 13C-isotope labelling for the facilitated NMR analysis of a complex dynamic chemical system. *Chem Commun (Camb)* **2011,** *47*, 12476-8.

56. Pomerantz, W. C.; Hadley, E. B.; Fry, C. G.; Gellman, S. H., In Situ Monitoring of Backbone Thioester Exchange by F-19 NMR. *Chembiochem* **2009,** *10*, 2177-2181.

57. Caraballo, R.; Dong, H.; Ribeiro, J. P.; Jimenez-Barbero, J.; Ramstrom, O., Direct STD NMR Identification of beta-Galactosidase Inhibitors from a Virtual Dynamic Hemithioacetal System. *Angew Chem Int Edit* **2010,** *49*, 589-593.

58. Lienard, B. M. R.; Hueting, R.; Lassaux, P.; Galleni, M.; Frere, J. M.; Schofield, C. J., Dynamic combinatorial mass spectrometry leads to metallo-beta-lactamase inhibitors. *J Med Chem* **2008,** *51*, 684-688.

59. Lienard, B. M.; Selevsek, N.; Oldham, N. J.; Schofield, C. J., Combined mass spectrometry and dynamic chemistry approach to identify metalloenzyme inhibitors. *Chemmedchem* **2007,** *2*, 175-9.

60. Lienard, B. M.; Huting, R.; Lassaux, P.; Galleni, M.; Frere, J. M.; Schofield, C. J., Dynamic combinatorial mass spectrometry leads to metallo-beta-lactamase inhibitors. *J Med*

Chem **2008,** *51*, 684-8.

61. Ludlow, R. F.; Otto, S., Two-vial, LC-MS identification of ephedrine receptors from a solution-phase dynamic combinatorial library of over 9000 components. *J Am Chem Soc* **2008,** *130*, 12218-+.

62. Clark, M. A.; Acharya, R. A.; Arico-Muendel, C. C.; Belyanskaya, S. L.; Benjamin, D. R.; Carlson, N. R.; Centrella, P. A.; Chiu, C. H.; Creaser, S. P.; Cuozzo, J. W.; Davie, C. P.; Ding, Y.; Franklin, G. J.; Franzen, K. D.; Gefter, M. L.; Hale, S. P.; Hansen, N. J. V.; Israel, D. I.; Jiang, J. W.; Kavarana, M. J.; Kelley, M. S.; Kollmann, C. S.; Li, F.; Lind, K.; Mataruse, S.; Medeiros, P. F.; Messer, J. A.; Myers, P.; O'Keefe, H.; Oliff, M. C.; Rise, C. E.; Satz, A. L.; Skinner, S. R.; Svendsen, J. L.; Tang, L. J.; van Vloten, K.; Wagner, R. W.; Yao, G.; Zhao, B. G.; Morgan, B. A., Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat Chem Biol* **2009,** *5*, 647-654.

63. Melkko, S.; Scheuermann, J.; Dumelin, C. E.; Neri, D., Encoded self-assembling chemical libraries. *Nat Biotechnol* **2004,** *22*, 568-574.

64. Brunschweiger, A.; Iqbal, J.; Umbach, F.; Scheiff, A. B.; Munkonda, M. N.; Sevigny, J.; Knowles, A. F.; Muller, C. E., Selective nucleoside triphosphate diphosphohydrolase-2 (NTPDase2) inhibitors: nucleotide mimetics derived from uridine-5'-carboxamide. *J Med Chem* **2008,** *51*, 4518-28.

65. Kogami, M.; Koketsu, M., An efficient method for the synthesis of selenium modified nucleosides: its application in the synthesis of Se-adenosyl-l-selenomethionine (SeAM). *Org & Biomol Chem* **2015,** *13*, 9405-9417.

66. Zameo, S.; Vauzeilles, B.; Beau, J.-M., Direct Composition Analysis of a Dynamic Library of Imines in an Aqueous Medium. *Eur J Org Chem* **2006,** *2006*, 5441-5444.

CHAPTER 4

The acceptor analogue GlcNAcβ1-(2-naphthyl): substrate or inhibitor of β-1,4-galactosyltransferase?

4.1. Glycosyltransferase assays

In order to explore the potential of GTs as molecular targets, a variety of bioassays have been developed.[1, 2] These assays allow the monitoring of GT activity during and after enzyme purification and can be applied for mechanistic studies of glycosylated transfer reaction, inhibition evaluation and high throughput screening (HTS) of GT inhibitors.

Functional GTs bioassays are employed with saturated substrate concentrations and depend on monitoring of depletion of substrate (both donor and substrate) or the formation of products (like glycosylated product or the secondary products from nucleotide diphosphates of reactions).

4.1.1. GT assays monitoring the primary glycosylation product

GT assays can be applied by monitoring the formation of primary glycosylation product. A variety of methods have been utilised to GT assays, like spectrophotometric methods and radiochemical methods, etc. Radiometric assay is one of the most extensively used GTs assays due to its high sensitivity and detection of low levels of enzyme. The design of the assay is based on the transfer of radio-labelled sugar nucleotide donor to an acceptor. The radio-labelled saccharide product can be quantified by scintillation counting method.[3] The radio-labelled sugar nucleotides derivatives are commercially available. The radiometric assay involving with radio-labelled donors and acceptors usually is non-continuous assay and after quenching the reaction, a variety of methods have been employed to remove the un-

turnover radio-labelled donors, such as electrophoresis, ion-exchange chromatography, TLC, size-exclusion chromatography for glycoproteins.[3-6] As the separation and wash steps are required, radiometric assay is difficult for HTS application. In order to overcome the difficulty, scintillation proximity assays (SPAs) have been developed for GTs assay. For example, a miniaturized scintillation proximity assay was utilized for the HTS screening of drug-like glycosidic acceptor inhibitors towards human FucTVII.[7] The used acceptor was modified with scintillat-impregnanted microspheres which could emit light when stimulated with radiolabelled molecules and the signal could be detected directly without separation or wash steps.

The non-natural substrates, like labelled donors or acceptors, can be utilized in functional GT assays. The labelled products can be detected by spectrophotometric methods, like UV and fluorescence measurement, and radiochemical methods (scintillation counting). (**Figure 70**) For example, the fluorescent moiety containing substrate, C6-4-nitrobenzo-2-oxa-1,3-diazole (NBD)-ceramide, and C6-NBD-glucosylceramide, were utilised to determine GlcTs and GalTs activities and a normal-phase HPLC was applied for its measurement.[8] The preparative synthesis is required for the fluorophore/chromophore labelled enzyme substrates and the substrate derivatives may hamper the glycosylated reactions. Thus, new methodology was developed to overcome these difficulties. α–1,3-galactosyltransferase activity was monitored by a fluorephore-assisted capillary electrophoresis (FACE) method.[9]

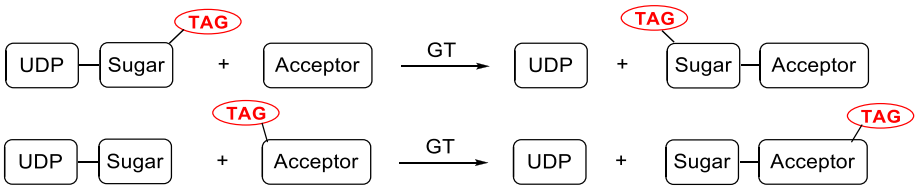


Figure 70 Non-natural substrate, like labelled donor or substrate: the enzyme activity can be determined depending on the product formation by spectrophotometric or fluorescent methods.

Mass spectrometry (MS) has been applied for GTs assay. MS can be used to directly determine both the formation of glycosylated product and substrate depletion. It can be applied for GTs inhibitor screening[10] and kinetic characteristic[11] and it also can be adapted for HTS[11].

4.1.2. GT assays monitoring the secondary glycosylation product

GT assays can be applied depending on the formation of secondary products of reactions, nucleotide diphosphates. Due to the presence of uracil moieties, nucleoside diphosphates are UV-active and the HPLC, capillary electrophoresis (CE), ion-exchange chromatography methods can be employed for GT assays by monitoring the formation of relating nucleoside diphosphate. [9, 12-15] (**Figure 71**)

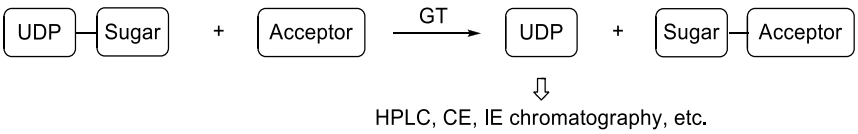


Figure 71 Native GTs reactions, the enzyme activity can be determined depending on the formation of NDP by HPLC, CE, IE exchange, etc.

Coupled enzyme assay are performed with natural GT substrates and they generally are performed in a multi-wells plate, making them amenable for high-throughput formats. As many GTs utilize sugar nucleotides as natural donors, producing UDP or UMP in GT reactions, these assay formats are universal and can be applied to many GTs. A continuous spectrophotometric assay was carried out for determining kinetic parameters for three GTs,

fucosyl-, sialyl-, N-acetylgucosaminyltransferases, by coupling enzyme product UDP to NADH oxidation *via* pyruvate kinase and lactate dehydrogenase. The reaction progress was monitored by absorbance at 340 nm.[16]

A phosphatase-coupled glycosyltransferase assay takes advantage of specific phosphatases which can be utilized in GT reactions to quantitatively release inorganic phosphatase from the leaving nucleoside diphosphates of GT reactions. The released inorganic phosphate group can be quantified by colorimetric detection with Malachite Green.[17, 18] (**Figure 72**) The amount of inorganic phosphate produced is proportional to the sugar transferred in GT reaction, hence the reactivity of GTs can be measured. This assay can be applied for high-throughput screening for inhibitor development. As this method is based on monitoring the formation of secondary product from nucleoside diphosphate, it is essential to deduct the enzymatic or hydrolysis of the sugar donor under assay conditions which may cause positive results.[17] Besides GTs, other enzyme, phosphatase, is utilized in this assay and the inhibitor candidates also might interfere the phosphatase or chemically react with other assay reagents to confound the assay readout by producing apparent inhibitory activities. Thus, the control assay is required to exclude the false positive results.

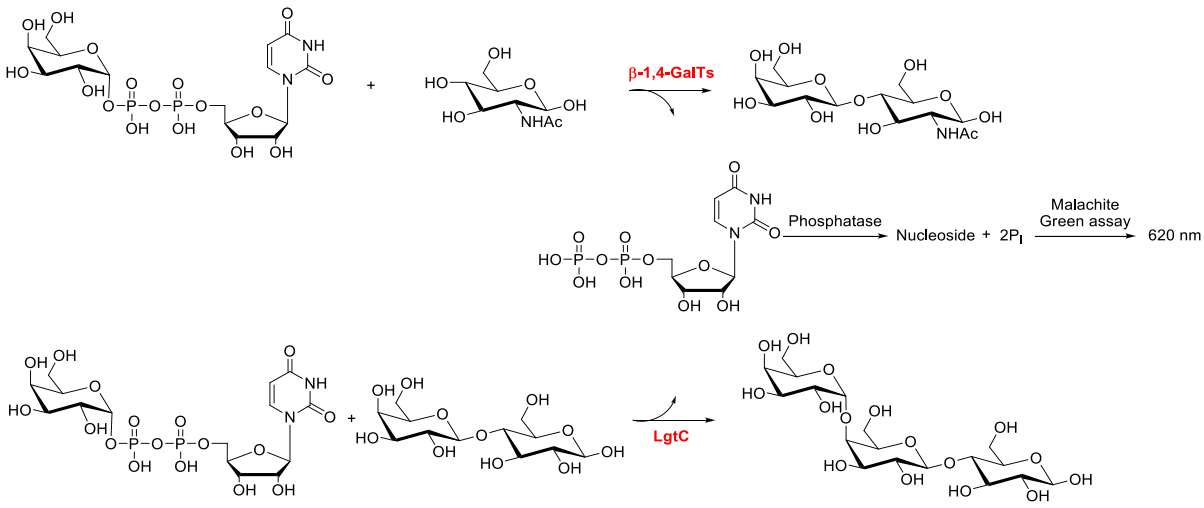


Figure 72 Principle of the phosphatase-coupled glycosyltransferase assay used in this study

Recently, innovative GTs assays have been developed. For example, a label-free GTs assay was designed to monitor enzyme activity by using an artificial chemosensor.[19] This chemosensor could selectively bind to pyrophosphate monoesters but not diester, leading to a significant increase of fluorescent signal. This approach was applied to two galactosyltransferases which can produce nucleotide diphosphate for identification of their inhibitors. Based on the principle, it can be broadly applied to many GTs that generate NDPs and can be adapted for HTS.

These assays all provide essential information about the biological activity of the used GTs and can be used for inhibitor validation and substrate specificity studies. The GT assays monitoring the formation of secondary product, like enzyme coupled assay, could be performed in the multi-wells plates. Thus, these assays are applicable for HTS assays. On the contrary, these assays consume large quantities of valuable reagents and enzymes when applied to large compound library screening. Advantage of these assays is that they can be applied to the GT reactions that produce nucleoside diphosphate. As the separation and wash steps are required, the GT assays monitoring the formation of primary glycosylation product, like radiometric assay, are usually difficult for HTS application. These assays are sensitive to detect small concentration changes in GT-catalysed reactions for monitoring the enzyme activities.

4.1.3. Other GT assays

Many other methods have been developed for GTs activity measurement. For example, an equivalent proton is released during the transfer of sugar moiety from sugar nucleotides to

acceptors. This change of pH in assay can be detected *via* using the standard pH indicator phenol red for monitoring enzyme activity and substrate specificity.[20] Immunological assays, enzyme-linked immunosorbant assay (ELISA) also have been developed for GTs assays.[21, 22] With highly specific antibodies or lectins, immunological assays can be applied for HTS.[23] Fluorescence-based ligand displacement assays that monitor ligand binding are an attractive technique for large inhibitor candidate library screening.[24-26] (**Figure 73**) The assays format focuses on monitoring the binding of donor towards enzyme, not the conversion. Thus, GTs acceptors are not required. The ligand displacement assays are suitable for HTS for identifying inhibitors due to its high sensitivity and simple manipulation.[27]

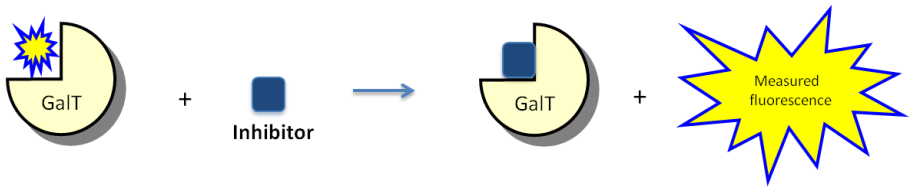


Figure 73 Principle of fluorescence-based ligand displacement assay

The Fluorescence Resonance Energy Transfer-based (FRET-based) assay by taking advange of the enzymatic breaking of the glycosidic linkage was developed for GTs.[28, 29] This assay is based on the distance dependent transfer of energy from a donor molecule to an acceptor molecule. The transfer of energy leads to a reduction in the donor's fluorescence intensity and an increase in the emission intensity of acceptor. As the donor and acceptor moiety are different, in this case FRET can be detected by the appearance of fluorescence of the acceptor or by quenching of donor fluorescence. (**Figure 74**)

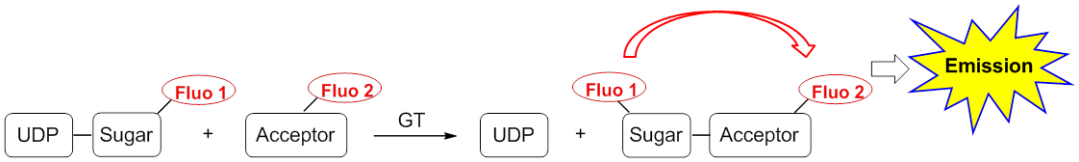


Figure 74 Principle of FRET-based assay

4.2. Objectives

For the discovery of new β-1,4-GalT inhibitor chemotypes *via* DCC experiment, a known β-1,4-GalT inhibitor was required as a positive control. GlcNAc β1−(2-naphthyl) **43** is a GlcNAc derivative with a naphthyl aglycone, which has been previously reported as an acceptor-based inhibitor of β-1,4-GalT.[30] (**Figure 75**) In the absence of a commercially available β-1,4-GalT inhibitor, GlcNAc derivative **43** was selected as the positive control due to its simple, two-step synthesis and reasonable inhibitory potency (72 % inhibition at 0.4 mM) It was reported that the hydrophobic properties of the naphthyl ring structure at the anomeric position may enable the substrate analogues to tightly bind to enzyme, forming potent inhibitors.[31]

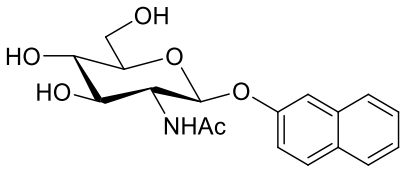


Figure 75 GlcNAc derivative GlcNAc β1-(2-naphthyl) **43**

In order to characterize the biological activity of compound **43**, a number of biochemical experiments were carried out. The unexpected results of these experiments, which differed

from the previously reported literature[30], are described in this chapter (**Figure 76**) In order to explain the observed discrepancies and unambiguously establish the activity of **43** towards β-1,4-GalT, a series of experiments was carried out, including the identification of the primary glycosylation product by LC/MS.

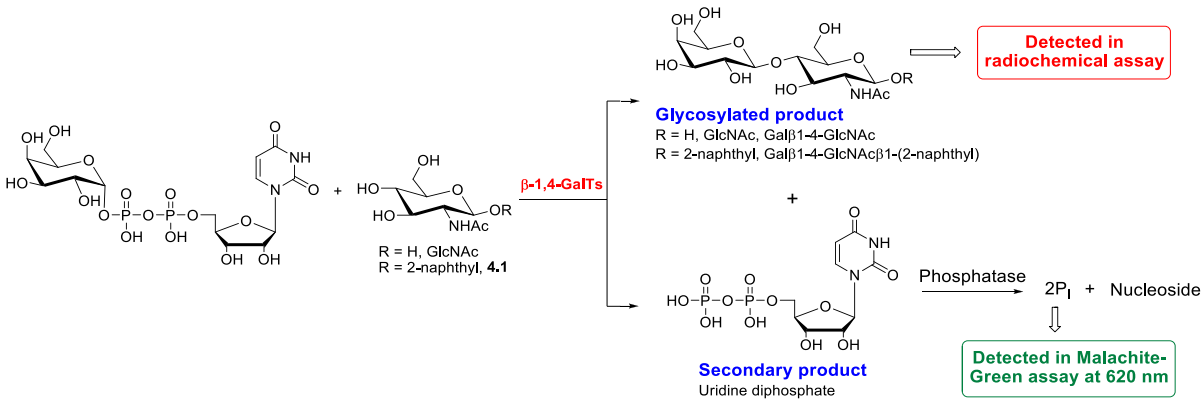


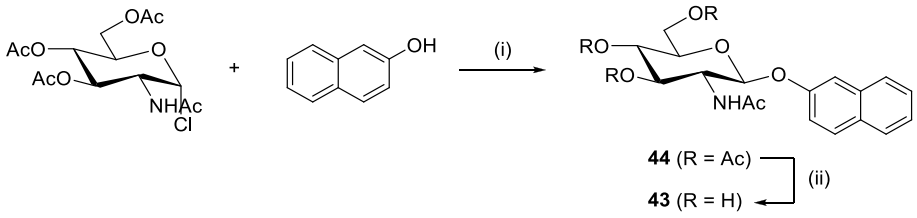
Figure 76 Discrepant results in different glycosyltranferases assay

4.3. Results and discussion

4.3.1. Synthesis of GlcNAc β1-(2-naphthyl)

The synthesis of GlcNAc β1–(2-naphthyl) was carried out according to the procedure reported[32] (**Scheme 18**). The fully acetylated naphthyl glycoside **44** was prepared from 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride by nucleophilic substitution with 2-naphthol under phase-transfer catalytic conditions, in 75 % yield. Direct

de-protection of compound **44** using 0.05 M sodium methoxide afforded the desired product 1-(2-naphthyl) 2-acetamido-2-deoxy-β-D-glucopyranoside **43**.



Scheme 18 Synthesis of GlcNAc β1-(2-naphthyl) **43**. Reagents and conditions: (i) tetrabutylammonium bromide, 1N NaOH, CH_2Cl_2, rt, 2h, 75%. (ii) NaOMe, MeOH/Toluene = 1/1, rt, 0.5h, 88%.

Results from the analytical characterisation of compound **43** by ^{1}H, ^{13}C and mass spectrometry were consistent with the data of previous literature.[33] The coupling constant between H-1 and H-2 showed 8.4 Hz (**Figure 77**). This *J* value is indicative of the *trans* orientation of H-1 and H-2, and hence the β-configuration of the aglycone at the anomeric position of **43**.

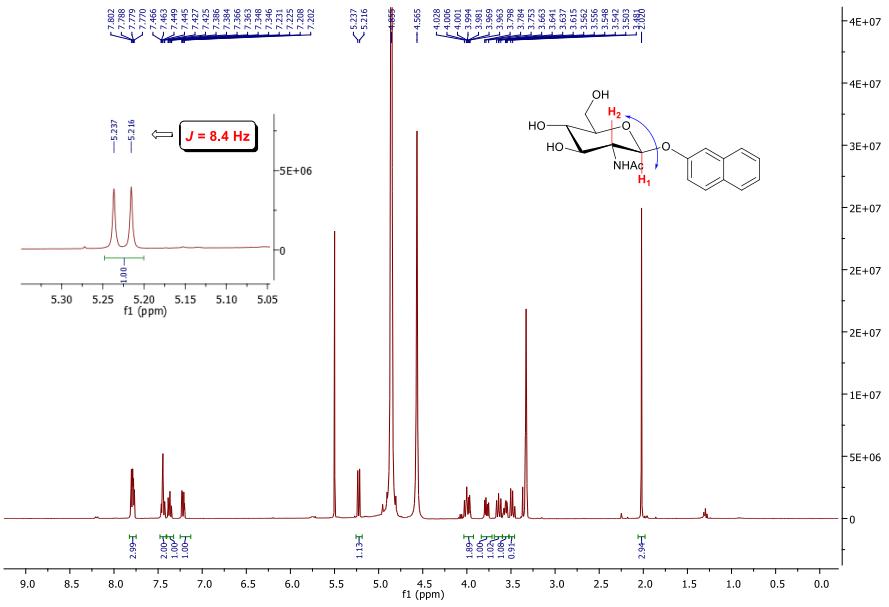


Figure 77 ^{1}H-NMR spectrum of compound **43** and the *J* value between H-1 and H-2

4.3.2. Bioassay for substrate activity assessment

With compound **43** in hand, the bioassay was carried out. It was reported to act as a potent inhibitor towards β-1,4-GalT.[30] In principle, this monosaccharide was an analogue of the natural donor GlcNAc, containing an additional naphthyl group in the aglycone moiety. It was reported that the 2-acetamido and the 4-hydroxyl group of GlcNAc were essential for binding of GlcNAc towards enzyme.[34] Therefore, compound **43** also may behave as a substrate towards the enzyme. Therefore, the assay was carried out for assessing the substrate activity of compound **43** towards bovine β-1,4-GalT. The biochemical phosphatase-coupled glycosyltransferase assay was used and the design of well map of assay was illustrated in **Figure 78**. For the substrate experiments, the standard natural acceptor substrate GlcNAc was replaced with derivative **43**. Experiments with either GlcNAc or the putative inhibitor **43** were conducted in parallel in the same microplate (triplicate for each). The maximal concentrations of GlcNAc and its derivative were not identical, 5 mM for GlcNAc and 1 mM for compound **43**. It was because that in the standard inhibition bioassay, GlcNAc was used as the natural substrate at the concentration of 5 mM towards β-1,4-GalT and inhibitor candidates were evaluated towards β-1,4-GalT at the concentration of 1 mM. As the donor substrate, UDP-galactose, contributes to the background of the assay by enzymatic and chemical hydrolysis to terminal inorganic phosphate, so the relevant control wells in each microplate (no acceptor, but otherwise identical conditions, carried out in duplicate) was included in the same microplate to account for non-specific hydrolysis.

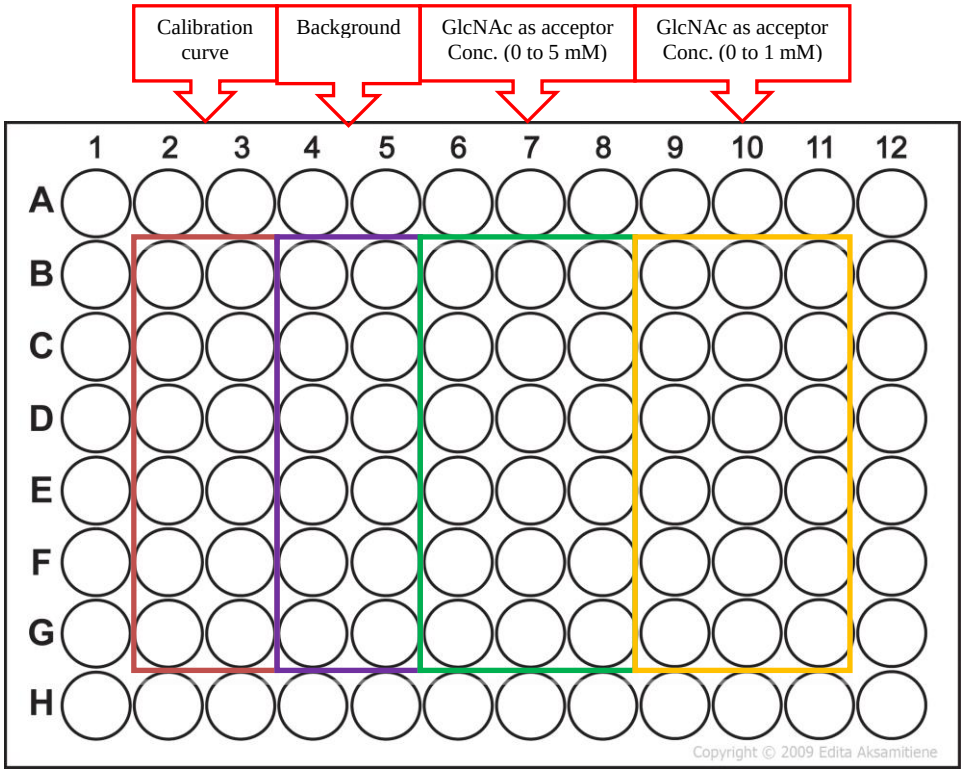
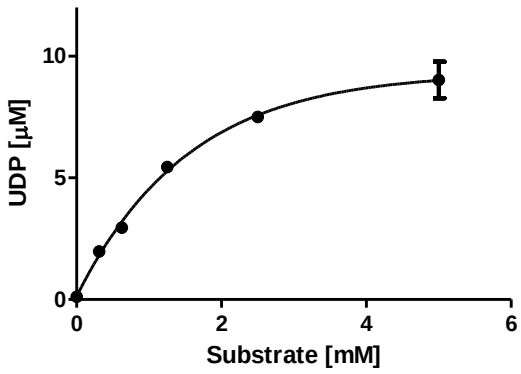


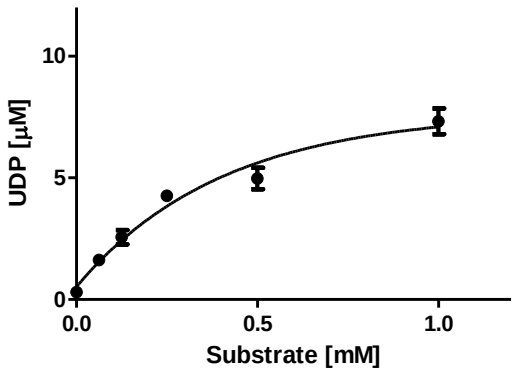
Figure 78 Well map design of substrate assay

In the presence of either GlcNAc or its GlcNAc derivative **43** as an acceptor substrate, significant β-1,4-GalT activity was observed, as measured by the formation of the secondary reaction product UDP. The observed β-1,4-GalT activity was dependent on the concentration of the respective acceptor (**Figure 79**). According to the graphs, in the case of GlcNAc derivative **43**, maximal activity was observed at lower acceptor concentrations than in the case of GlcNAc, which suggested that **43** may actually be a better acceptor substrate than GlcNAc. This result that compound **43** acted as substrate towards β-1,4-GalT was unexpected as it was discrepancy against the literature result. In order to confirm this discrepant result, control assays were then carried out.

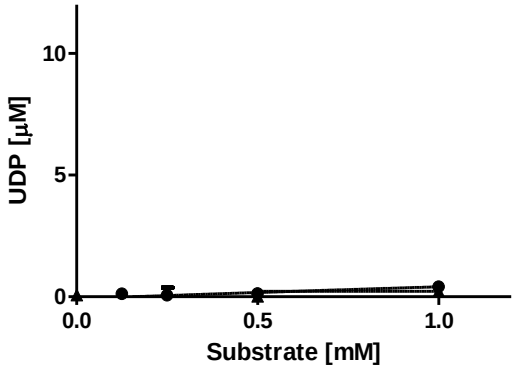
Figure 79 Acceptor substrate assays with recombinant β-1,4-GalT in the absence of BSA.

(a) β-1,4-GalT-catalysed conversion of UDP-Gal donor at different concentrations of GlcNAc (left) or compound **43** (right) as the acceptor

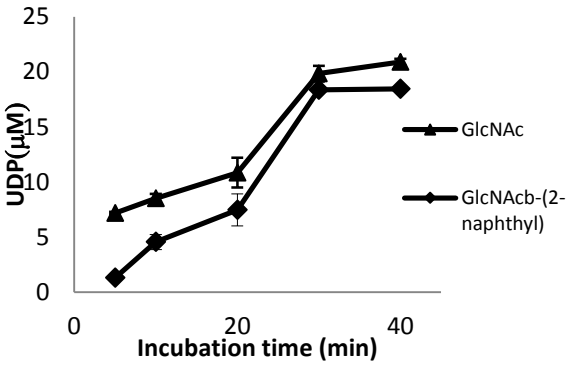




(b) Control experiment with compound **43** (no β-1,4-GalT)



(c) Time-dependent, β-1,4-GalT catalysed conversion of UDP-Gal donor with GlcNAc or GlcNAc □□-(2-naphthyl) **43** as the acceptor



Conditions: β-1,4-GalT, acceptor (GlcNAc, 0 to 5 mM/ **43** 0 to 1 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (1 mg mL^{-1}), calf-intestinal phosphatase (10 U mL -1), DMSO (10%) and buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 °C with shaking for required time (5 to 40 min). The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments.

The tested derivative **43** may chemically react with assay reagents to confound the assay readout by producing apparent biological activity. Therefore, a separate control experiment was carried out to check this mode of assay interference of compound **43**. For the control

assay, the assay condition was identical to the standard substrate assay except the addition of enzyme (**Figure 80**). As illustrated in **Figure 79b**, GlcNAc derivative **43** itself did not interfere with the colorimetric readout of the assay as evidenced by the lack of signal in the control experiment with increased concentration of **43**.

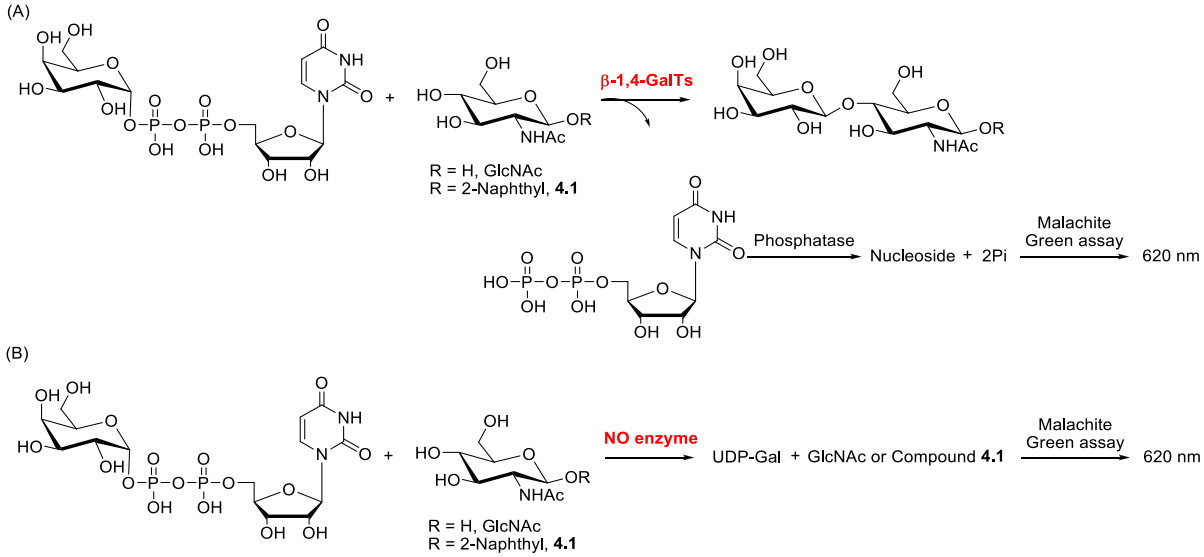


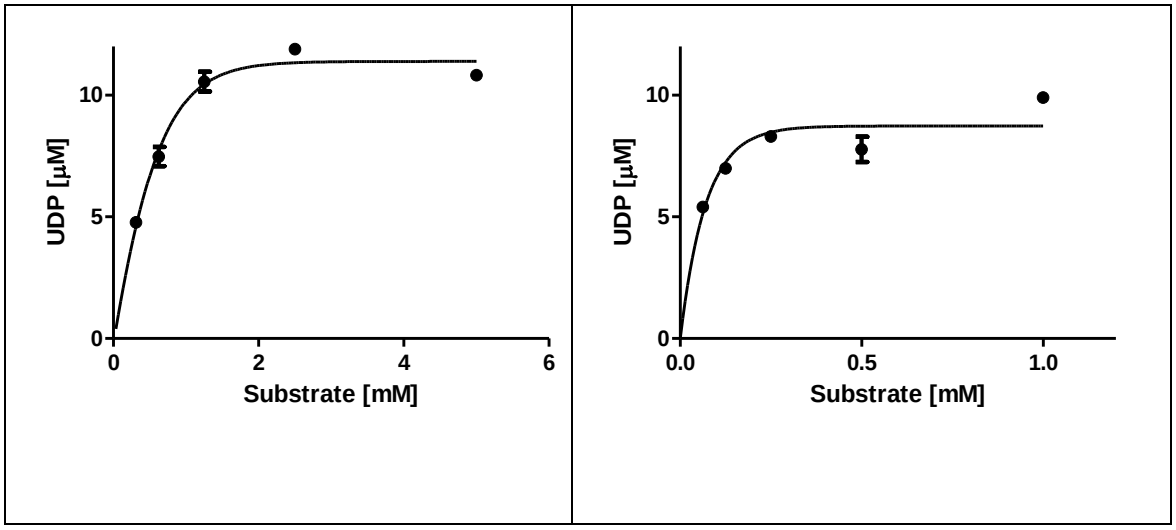
Figure 80 (A) Standard bioassay for assessing substrate activity of compound **43**. (B) Control assay for checking the reaction between compound **43** and other reagents in the assay.

In order to investigate the time dependency of the β-1,4-GalT reaction under these conditions, a single concentration of GlcNAc (5 mM) or its derivative **43** (1 mM) were utilized in the assay with 5 to 40 min incubation time at 30 ºC. A clear time-dependency of turnover was observed with both compounds (**Figure 79c**). By increasing incubation time, the formation of secondary product UDP increased in the presence of both GlcNAc and **43**, with a maximal turnover after 30 min. According to these results, it was strongly suggested that GlcNAc derivative **43**, just like its parent compound GlcNAc, was recognized as an acceptor substrate towards β-1,4-GalT.

In order to understand this discrepancy with previous literature reports, additional experiments were carried out. In the literature precedent, naphthyl-containing GlcNAc

derivative **3** and its analogues were evaluated in a radiochemical glycosyltranferase assay.[30] The phosphatase-coupled glycosyltransferase assay used in our present study is carried out by measuring the formation of secondary reaction product UDP. In this assay, the coupled phosphatase is utilized to hydrolyse inorganic phosphate quantitatively from UDP and the free phosphate can be quantified by phosphatase-coupled assay. (**Figure 80A**) In contrast, the radiochemical assay used in the original report is based on the transfer of radio-labelled sugar nucleotide donor to an acceptor. The radio-labelled saccharide product can be quantified by scintillation counting method. The enzyme used in the radiochemical assay of previous literature was obtained commercially from Sigma supplier, whereas our phosphatase-coupled assays were all carried out with a batch of bovine β-1,4-GalT recombinantly expressed in our own laboratory. Also, their radiochemical assay was carried out in the presence of 12.5 mg/mL bovine serum albumin (BSA), which was the main different reagents used. To assess the effect of BSA on the recognition of GlcNAc derivative **43**, the substrate activity bioassay was repeated in the presence of BSA (12.5 mg/mL). The microplate was incubated for 20 min and absorbance was recorded at 620 nm. When the formation of UDP was plotted against the concentration of acceptor (both GlcNAc and compound **43**), very similar reaction profiles between the two acceptor substrates were observed once again (**Figure 81**). Compared with the assay in the absence of BSA, the turnover of transfer reaction in the presence of BSA was improved slightly, and maximum turnover was observed with the concentration of GlcNAc (1.25 mM) and **43** (0.25 mM)

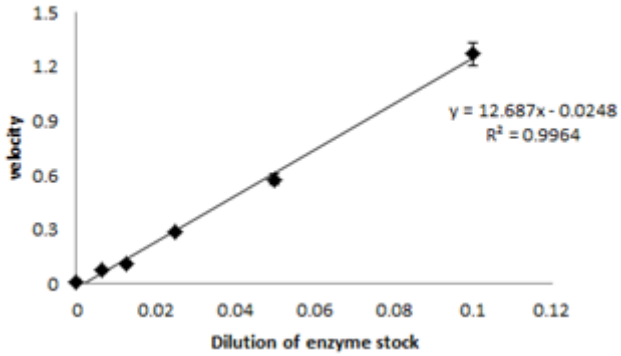
Figure 81 Acceptor substrate assays with recombinant β-1,4-GalT in the presence of BSA with GlcNAc (left) or compound **43** (right) as the acceptor.



Conditions: β-1,4-GalT, GlcNAc (0-5 mM) or **43** (0-1 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), chicken egg-white lysozyme (1 mg/mL), calf-intestinal phosphatase (10 U/mL), bovine serum albumin (1.25 mg/mL), DMSO (10%) in buffer (13 mM HEPES, 50 mM KCl, pH 7.0) were incubated in a 96-well plate at 30 °C with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All experiments were carried out in triplicate. Bars indicate mean values ± S.D.

In the previously reported literature, compound **43** was assessed with a commercially available batch of enzyme.[30] As the information provided by the supplier about this β-1,4-GalT batch was limited, it was uncertain whether it was identical to our own recombinant batch. Therefore, it was decided to repeat the experiments with a batch of commercial β-1,4-GalT from the same supplier. Firstly, the activity of this batch of β-1,4-GalT was determined after it was obtained. The rate of transfer reaction was plotted against the dilution of enzyme stock, with an activity of 12.687 mU/mL (**Figure 82**).

Figure 82 Enzyme activity measurement assay for commercial enzyme

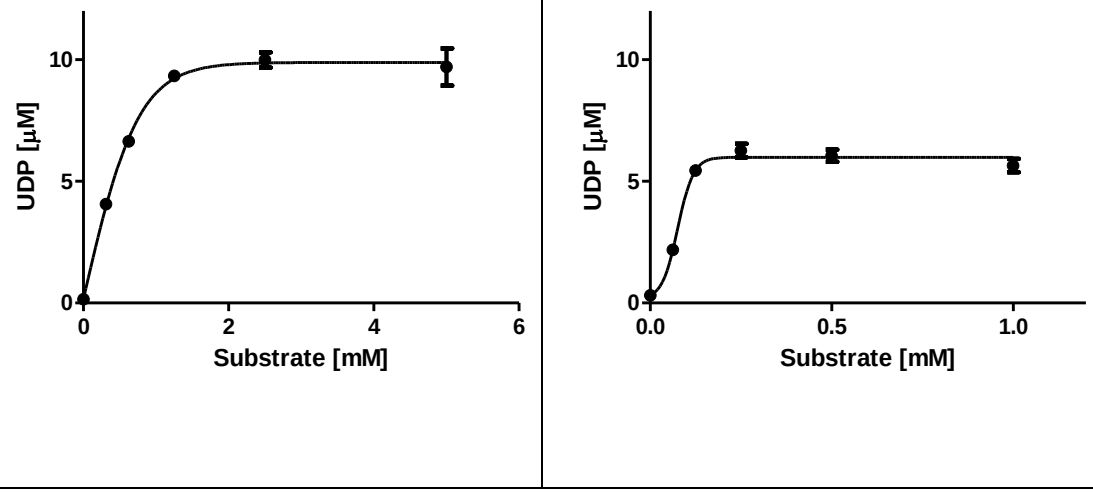


Conditions: β-1,4-GalTs (diluted to the required concentrations), acceptor (GlcNAc, 5 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), Chicken egg-white lysozyme (CEL, 1 mg/mL), calf-intestinal alkaline phosphate (CIP, 10 U/mL), DMSO (10%) buffer (13 mM HEPES, pH = 7.0, 50 mM KCl) were incubated on a 96-well plate at 30 °C with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. All concentrations are final concentrations. Bars indicate mean values ± S.D. of triplicate experiments.

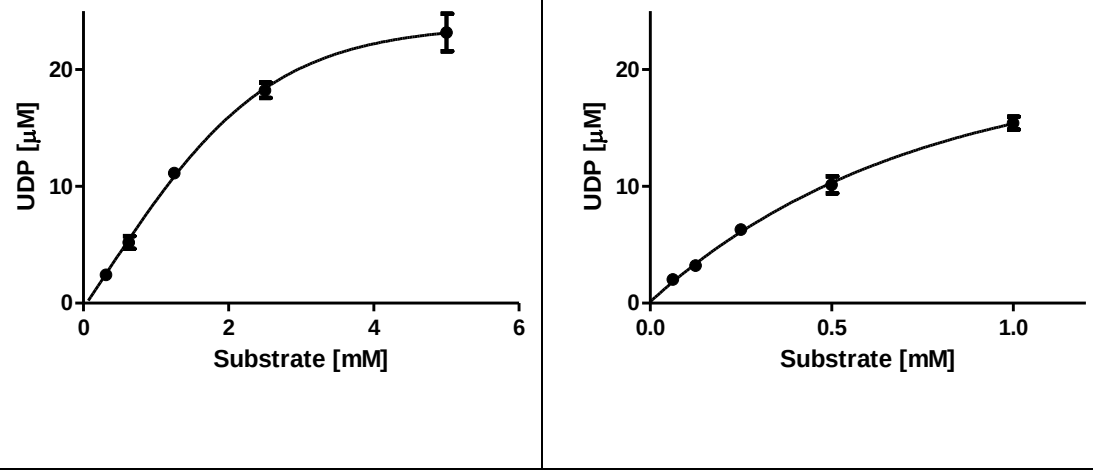
Next, the substrate assays with either GlcNAc or **43** as acceptor was repeated in the absence or presence of BSA and the results are shown in **Figure 83**. The results were nearly identical. Both GlcNAc and the derivative **43** acted as substrates towards β-1,4-GalT, both in the absence and presence of BSA. Compared with the assay of recombinantly expressed enzyme, the turnover of transfer reaction with GlcNAc derivative **43** was lower than that with GlcNAc and in the presence of BSA, enzyme activity was increased, with improved turnover, suggesting that BSA could facilitate the catalytic reaction.

Figure 83 Acceptor substrate assays with commercial β-1,4-GalT in the absence or presence of BSA.

(a) β-1,4-GalT-catalysed conversion of UDP-Gal donor at different concentrations of GlcNAc (left) or compound **43** (right) as the acceptor, in the absence of BSA.



(b) β-1,4-GalT-catalysed conversion of UDP-Gal donor at different concentrations of GlcNAc (left) or compound **43** (right) as the acceptor, in the presence of BSA.



Conditions: (a) β-1,4-GalT, GlcNAc (0-5 mM) or **43** (0-1 mM), UDP-Gal donor (28 μM), $MnCl_2$ (5 mM), chicken egg-white lysozyme (1 mg/mL), calf-intestinal phosphatase (10 U/mL), bovine serum albumin (1.25 mg/mL), DMSO (10%) in buffer (13 mM HEPES, 50 mM KCl, pH 7.0) were incubated in a 96-well plate at 30 ^{o}C with shaking for 20 min. The reaction was stopped by the addition of malachite reagents, and the absorbance was recorded at 620 nm after 30 min. (b) Conditions as in (a) but with addition of bovine serum albumin (1.25 mg/mL). All experiments were carried out in triplicate. Bars indicate mean values ± S.D.

4.3.3. Analysis of phosphatase-coupled assays results

GlcNAc derivative **43** has been reported to be inactive as a substrate, but active as an inhibitor towards bovine β-1,4-GalT.[30, 36] However, in our phosphatase-coupled assays substrate assay, this compound behaved as an acceptor substrate of β-1,4-GalT with a similar

profile to the natural GlcNAc acceptor. The spectroscopic characterisation of GlcNAc derivative **43** confirmed that this compound was identical to the one used in the previous report, ruling out the possibility of mistaken identity of the compound. In the previous report, compound **43** was tested in the radiochemical glycosyltransferase assay. The activity of GlcNAc derivative **43** was evaluated according to the direct detection of production of galactosylated reaction product. Our phosphatase-coupled glycosyltransferase assay measures the formation of the secondary reaction product, UDP-galactose, of enzyme transfer reaction. Both the radiochemical GT assay and phosphatase-coupled assays are known to be amongst the most extensively utilized GT assays. In theory, there should be no significant deviation between the results of these two assays. Different from our protocol, the reagents BSA and commercial β-1,4-GalTs were employed in the previous study. As the results of our assays with either recombinantly-expressed enzyme or commercially available enzyme in the absence or presence of BSA did not show significant difference, the difference of these data for compound **43** should not originate from the reagent BSA and commercial β-1,4-GalTs. The presence of GlcNAc β1-(2-naphthyl) **4.3** led to a significantly enhanced signal. Also, relative control experiments were carried out either in the same or a different microplate to eliminate the possibility of false positive results. All the assays were repeated three times and the results were identical. These results of experiments suggested the behaviour of **43** as a substrate in the phosphatase-coupled glycosyltransferase assay. The assay which detecting the formation of primary glycosylation product was then carried out to support the result.

4.3.4. Detection of the primary glycosylation product

In our bioassay, the glycosylated product, Galβ1-4-GlcNAcβ1-(2-naphthyl), was not observed directly and only the production of UDP could not give enough evidence to explain the discrepancy with the previous literature. Therefore, experiments were carried out to directly detect the putative product resulting from glycosylation of GlcNAc derivative. As the starting material of the transfer reaction, compound **43**, UDP-Gal, as well as the target product Galβ1-4-GlcNAcβ1-(2-naphthyl) **45** are all UV active molecules, LC/MS allows us to monitor the production of galactosylated saccharide product (**Figure 84**).

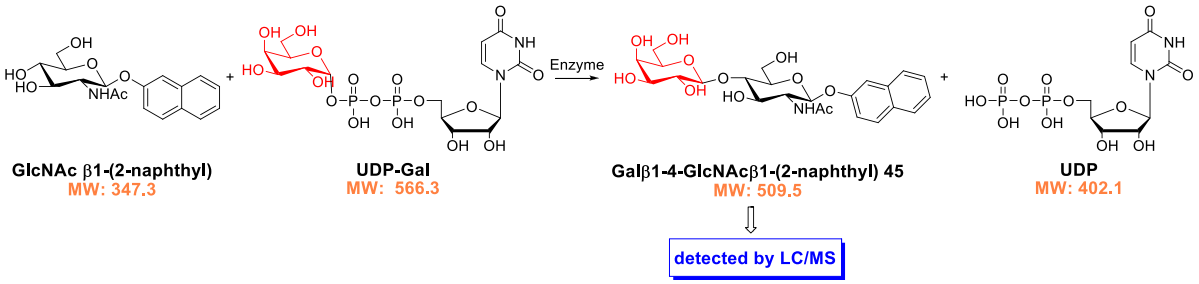


Figure 84 Different reaction species occurring during galactosylation of GlcNAcβ1-(2-naphthyl) and their molecular weights

Prior to the HPLC experiments, the UV absorbance of compound **43** was measured (**Figure 85**). The maximum absorbance of compound **43** is located at 220 nm. Since the UV active moiety of **43** and its galacosylated product **45** were the same, it was reasoned that the maximum UV absorbance wavelength would be identical. This allowed us to monitor the chromatographic elution from LC/MS at, or close, to the maximum absorbance exerted by these two compounds.

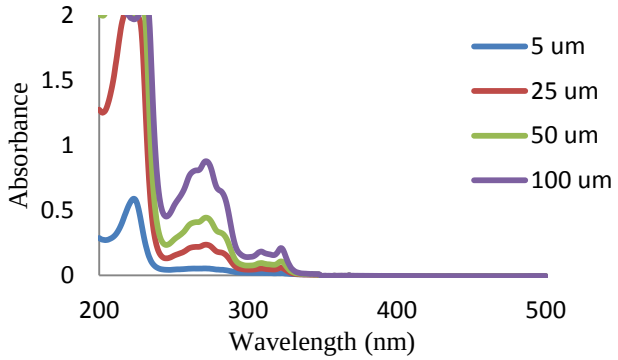


Figure 85 UV absorbance measurement of compound **43** in different concentration

The development of an LC/MS-based galactosyltransferase assay requires optimisation of both the enzymatic reactions and HPLC elution. It was anticipated that the highly polar compounds (UDP-Gal and by-product UDP) would be eluted quickly from the reverse-phase column. However, the separation of GlcNAcβ1-(2-naphthyl) **43** and the corresponding galactosylated product **45** would be more challenging because of their comparable polarity, which might require optimisation of separation conditions. A robust method was therefore needed to separate and identify the production of product **45**. We initially selected the HPLC system with a detection set at 214 nm and coupled with XDB-C8 column (5 μm, 4.6 × 150 mm) with a flow rate of 0.5 mL/min. The mobile phase comprised a mixture of water with 0.01 % formic acid (buffer A) and acetonitrile (buffer B) and the utilized gradient was illustrated in **Table 16**.

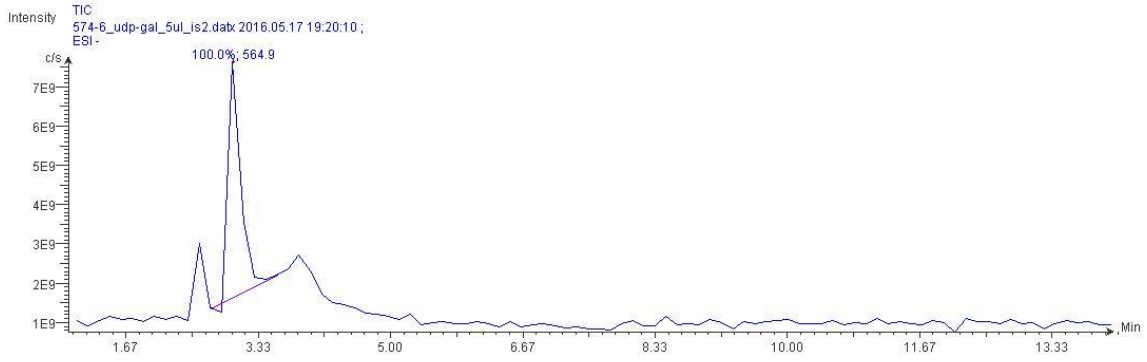
Table 16 Gradient utilised for separation of UDP-Gal and compound **43**

Gradient step	Time	Buffer A(%)	Buffer B(%)
1	1	90	10
2	13	90 to 20	10 to 80
3	14	20	80
4	15	95	5

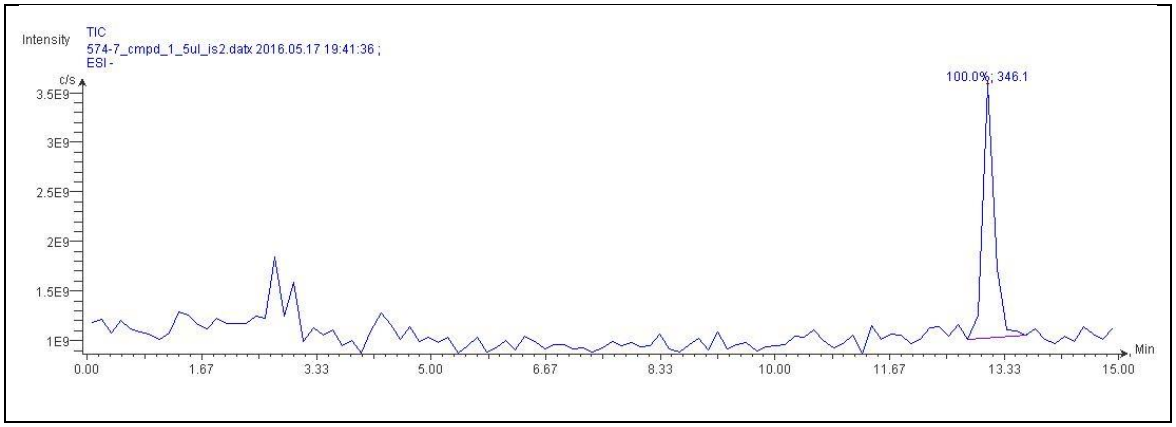
This elution system was utilized for UDP-Gal and **43** respectively (**Figure 86**). The retention time (tR) were 3.01 min and 13.1 min respectively. After achieving satisfactory peak resolution, the transfer reaction with β-1,4-GalT was carried out and analysed by this method. In order to simplify the assay system, in the initial experiments a mixture of enzyme, natural donor UDP-Gal and GlcNAcβ1-(2-naphthyl) **43** were incubated for 1h at 30 ^{o}C and this mixture was utilized for the analysis of LC/MS.

Figure 86 Chromatograms of UDP-Gal, compound **43** respectively. LC/MS performed on an Agilent 1200 series using XDB-C8 column (5 µm, 4.6 × 150 mm) with 0.5 mL/min flow rate with the detection at 214 nm. The mobile phase comprised a mixture of water (0.01% formic acid) and acetonitrile.

(a) Chromatogram of UDP-Gal (m/z 564.9, $[M-H]^-$), retention time = 3.01 min

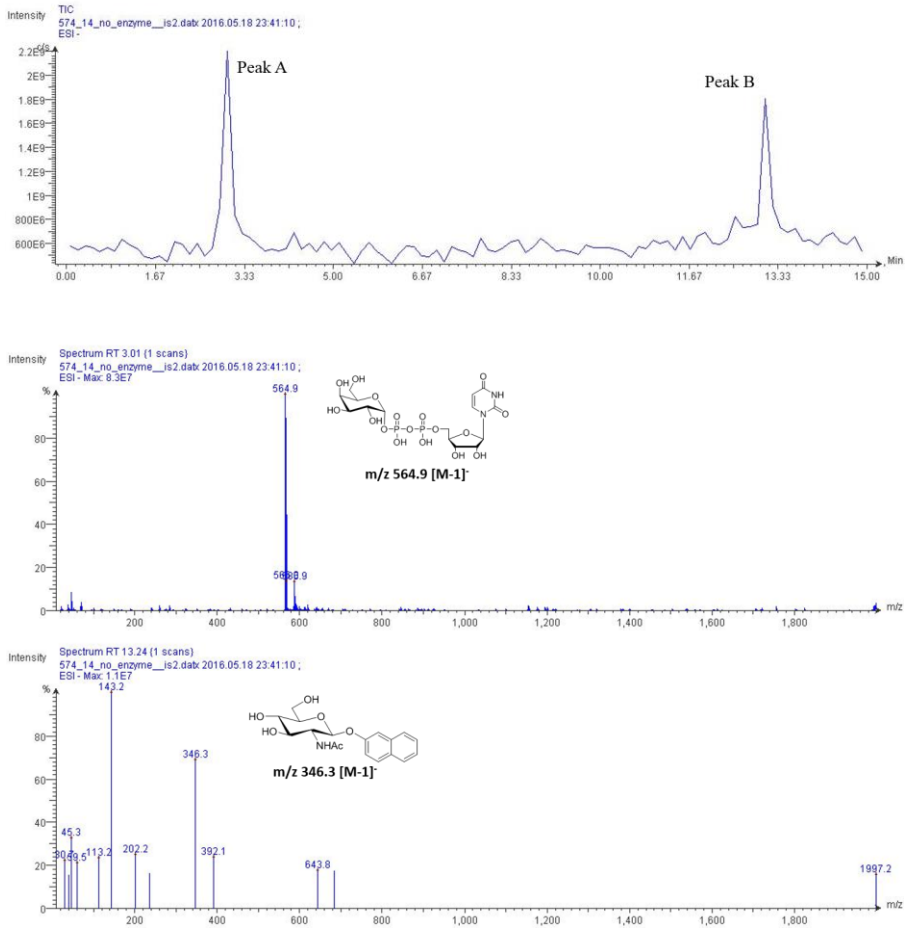


(b) Chromatogram of **43** (m/z 346.1 $[M-H]^-$), retention time = 13.1 min



The initial assay was carried out with UDP-Gal (500 μM), compound **43** (500 μM) and β−1,4-GalT (20 μL) in HEPES buffer (pH = 7.0). The mixture was incubated for 1h at 30 ^{o}C. The inactive protein might be precipitated from the buffer in the process of incubation. It would cause HPLC system block if the reaction solution was directly utilized into the equipment. Thus, the reaction was quenched by the addition of the same volume of methanol. After thorough mixing, the mixture was centrifuged for 15 min at 1000 rpm. The supernatant was removed for LC/MS analysis. However, under these conditions, no product peak for galactosylated product **45** (m/z 508 $[M-H]^-$) was detected (**Figure 87**).

Figure 87 Chromatogram of GalTs catalyzed reaction assay. 30 µL of reaction mixture was injected. LC/MS performed on an Agilent 1200 series using XDB-C8 column (5 µm, 4.6 × 150 mm) with 0.5 mL/min flow rate with the detection at 214 nm. The mobile phase comprised a mixture of water (0.01% formic acid) and acetonitrile.

Chromatogram of transfer reaction. Peak A: UDP-Gal (m/z 564.9, $[M-H]^-$), retention time = 3.01 min; Peak B: compound **43** (m/z 346.2 $[M-H]^-$), retention time = 13.1 min.


Reviewing the conditions for the enzymatic reaction and the LC/MS experiment, it was hypothesized that several factors may have prevented the formation and/or detection of the galactosylated product. Firstly, the glycosylated product was too little to be detected because of the use of limited amount of GalT. Secondly, as discussed before, the separation of compound **43** and its corresponding galactosylated product **45** was difficult due to their similar polarity. Therefore, insufficient separation may have led to the overlap of their peaks under these LC/MS conditions. Also, the irregular baseline may influence the appearance of

product peaks. In order to prevent the occurrence of these interferences, both the assay and LC/MS conditions were optimised. The amount of enzyme in the assay was increased in order to promote the formation of glycosylated product. Then, the HPLC elution system was modified as well. The buffer B (acetonitrile) was replaced by MeOH and a new gradient was developed. (**Table 17**)

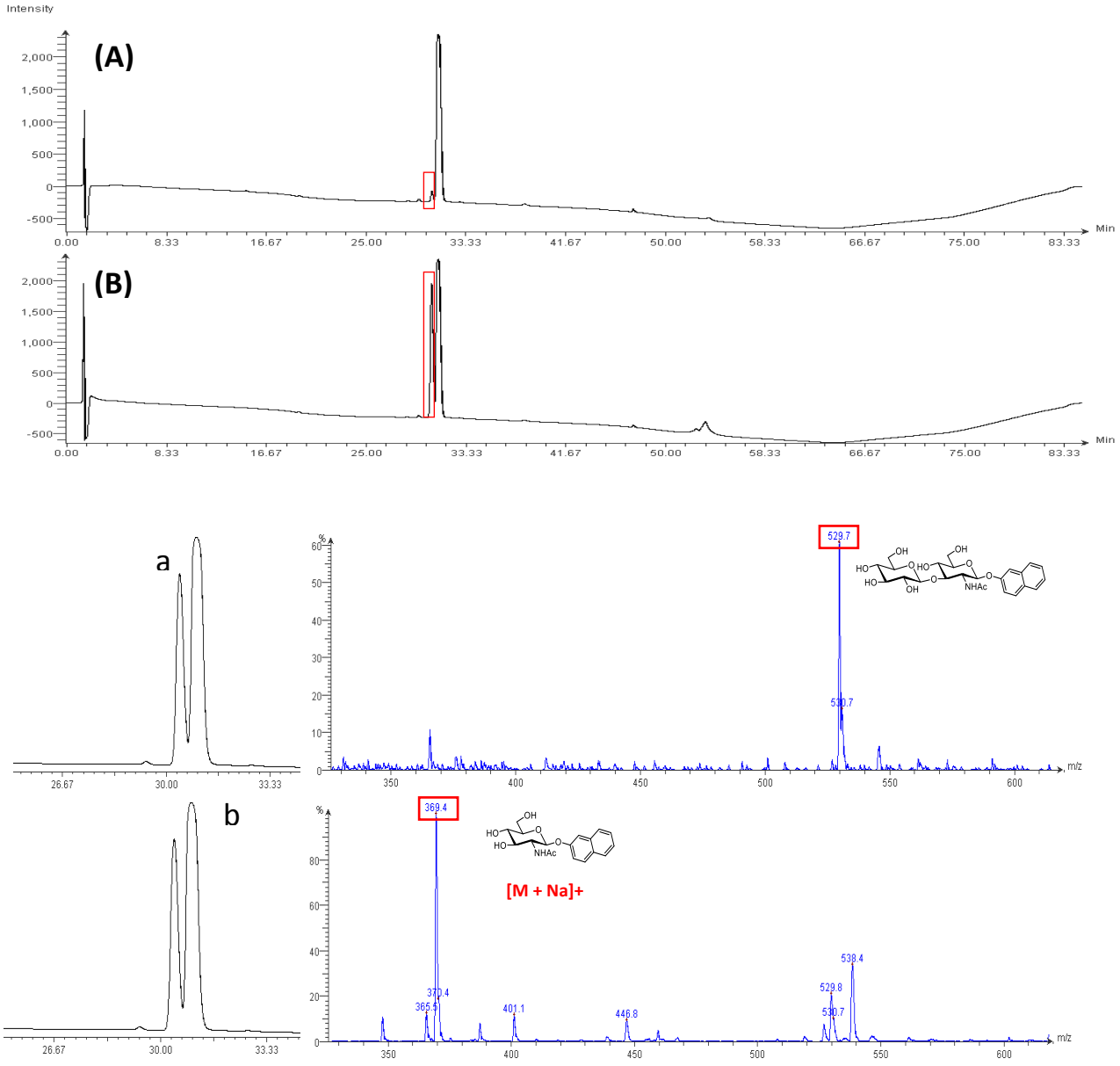
Table 17 Gradient B utilised for separation of UDP-Gal and compound **43**

Entry	Time (min)	Buffer A (%)	Buffer B (%)
1	1	90	10
2	10	90 to 70	10 to 30
3	20	70 to 50	30 to 50
4	30	50 to 30	50 to 70
5	40	30 to 20	70 to 80
6	70	20	80
7	80	20 to 90	80 to 10
8	85	90	10

Next trial of the assay was tested by using 50 μM enzyme. After 1h incubation, the mixture was used for LC/MS analysis. In addition to the peak for compound **43**, a second small peak appeared at slightly shorter retention time (**Figure 88**). As expected, the use of β-1,4-GalT with 200 μM resulted in increased formation of this peak (**Figure 90, assay B**). This pattern suggested that the transfer reaction did indeed occur, and that compound **43** acted as a substrate towards β-1,4-GalT. However, the peak **a** (m/z 529.7) did not fit the expected molecular mass of the glycosylated product, even though the correct molecular mass of the glycosylated product **45** had been obtained in previous experiments (m/z 508 $[M-H]^-$, **Figure 89**), which testified that the setup of our LC/MS experiment worked well. The experimental condition was identical to the previous experiments and the LC/MS experiment was repeated three times, ruling out the possibility of any technical errors. Compared with the chromatogram of previous experiment, peak **a** (m/z 529.7) exhibited identical retention time

(t_R = 31.3 min). Also, the assay only contained **43**, UDP-Gal and enzyme. These simple constituents could not possibly produce any other by-product. It was speculated that the assay mixture was stored in freezer for more than two week before utilized in LC/MS and the pH of the mixture solution might change, influencing the MS results. The mass of peak **a** was MW plus 20.5. It is conceivable that the pH change might influence the steric structure of **45**, contributing to different ion adduct, while we have no definitive explanation now for which species giving this unexpected mass.

Figure 88 LC/MS analysis showing the formation of a reaction product when compound **43** is used as the sole acceptor substrate for GalT.



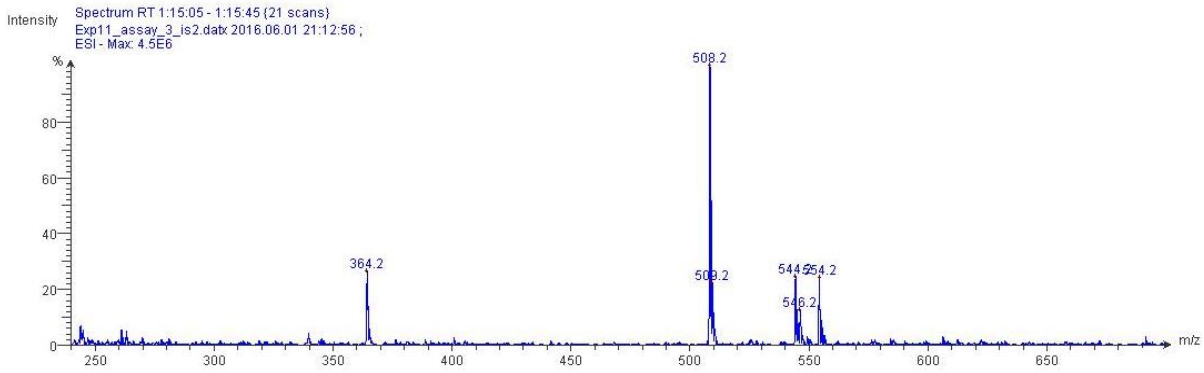
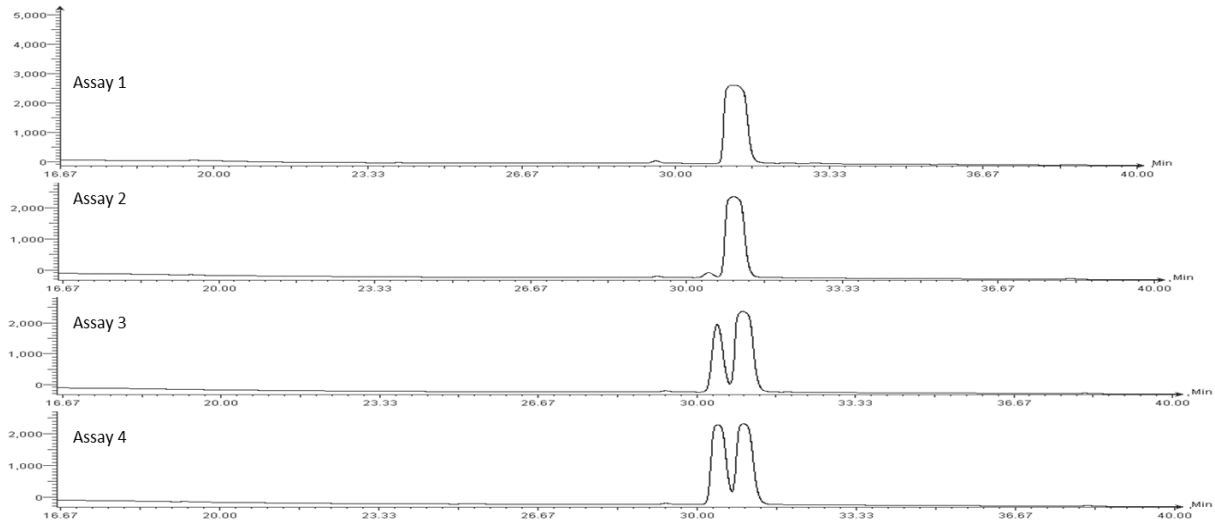


Figure 89 Mass spectra of glycosylated product **45** with correct molecule mass

Since phosphatase was utilized in the phosphatase-coupled assay protocol, we then investigated the assay condition with extra phosphatase added. In order to compare with the previous result, the assay with 50 μM β-1,4-GalT, UDP-Gal and compound **43** were incubated in the presence of 20 μM phosphatase. Compared with other assays, a significantly increased yield of product was obtained following the addition of phosphatase. (**Figure 90**, **Assay 4**).

Figure 90 Effect of the β-1,4-GalTs and phosphatase on the production of glycosylated product. (1) Assay conducted with compound **43** only. (2) Assay conducted with UDP-Gal, cmpd **43**, β-1,4-GalTs (50 μL). (3) Assay conducted with UDP-Gal, cmpd **43**, β-1,4-GalTs (200 μL). (4) Assay conducted with cmpd **43**, β-1,4-GalTs (50 μL), phosphatase (20 μL).



The structure of compound **43** is similar to that of the natural acceptor GlcNAc. It is speculated to bind in the same binding site as the substrate. According to the glycosyl transfer reaction mechanism, upon binding a donor and metal ion to enzyme, the conformation of enzyme could change from open to close conformation, creating substrate binding site. Then **43** binds with enzyme in the substrate site due to its better affinity towards enzyme. At some point, this reaction would reach the equilibrium. Thus, GlcNAcβ1-(2-naphthyl) **43** acts as an inhibitor towards enzyme in the radiometric assay which assesses molecule activity by detecting the formation of product. (**Figure 91**) However, the consumption of UDP was detected in our assay and β-1,4-GalT activity was dependent on the concentration of **43**. In the phosphatase-coupled glycosyltransferase assay, it is hypothesised that the use of phosphatase in the assay facilitates glycosylated reaction, which leads to the discrepant result. Phosphatase is used to hydrolyse the secondary product of glycosyl transfer reaction, UDP, into inorganic phosphate which was quantified by the absorbance measurement. In the catalytic cycle, after the binding of substrate towards enzyme, the sugar will transfer to the substrate and then the product dissacharide as well as the by-product, uridine di-phosphate (UDP), are ejected. Due to the presence of phosphatase, UDP is hydrolysed by phosphatase into inorganic phosphate, which is irreversible. As UDP is also reported as a GalTs inhibitor, the hydrolysis of UDP can maintain the enzyme activity and facilitate the formation of glycosylated products. Therefore, the glycosylated reaction is facilitated towards the production of inorganic phosphate and disaccharide product by the phosphatase and compound **43** acted as a substrate towards the enzyme. In the LC/MS experiment, a significantly increased yield of product was obtained due to the addition of phosphatase. Thus, this experimental result supported our explanation.

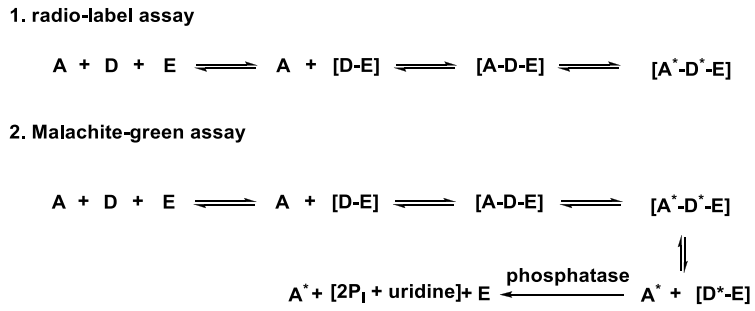


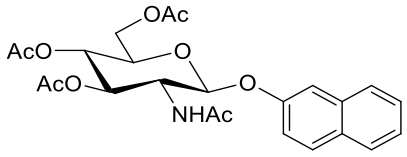
Figure 91 Proposed explanation for the differential results for substrate and inhibitor activity of compound **43**. A: GTs acceptors; D: GTs donors; E: GTs; A^*: glycosylated acceptors; D^*: nucleoside diphosphate (NDP).

4.4. Conclusion

It was found that GlcNAc β1-(2-naphthyl) **43** could behave as an acceptor substrate towards β-1,4-GalTs from bovine milk. This finding is in contrast with previously reported results that compound **43** acts as an inhibitor, but not an acceptor substrate of this enzyme. This results. According to the assay results, a hypothesis may be the presence of phosphatase in phosphatase-coupled assay. The phosphatase facilitates GT reactions by the hydrolysis of UDP, leading to the differential result from that obtained via radio-chemical assay. This observation has important implications for the application of compound **43** as a tool compound in glycobiology and glycobiochemistry. The result also emphasized the reagents utilized in assays that might interfere the detection of molecules' activities.

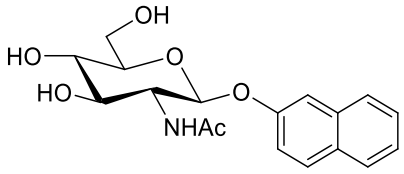
4.5. Experimental section

General. All reagents were obtained commercially and used as received unless stated otherwise. GlcNAc derivative **43** was synthesized as previously reported[30] and characterised by ^{1}H- and ^{13}C-NMR spectroscopy and mass spectrometry. Thin-layer chromatography (TLC) was performed on pre-coated aluminium plates (Silica Gel 60 F_{254}, Merck) and compounds were visualised by exposure to UV light (254 and 280nm). Preparative chromatography was carried out on silica gel 60 (pore size 60 Å, 230-400 mesh, Sigma-Aldrich) at normal pressure. NMR spectra were recorded at 298K on a Bruker Avance DRX 400 spectrometer (400 MHz for ^{1}H, 100 MHz for ^{13}C). Chemical shifts (δ) are reported in ppm (parts per million). Coupling constants (J) are reported in Hz.



2-Naphthyl 3,4,6-tri-*O*-acetyl 2-acetamido-2-deoxy-β-D-glucopyranoside (44) To an aqueous solution of sodium hydroxide (1M, 2 mL), a solution of 3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (400 mg, 1 equiv.), 2-naphthol (319 mg, 2 equiv.) and tetra-*n*-butylammonium bromide (354 mg, 1 equiv.) in methylene chloride (2 mL) was added. The resulting two-phase system was stirred for 2h at rt. The mixture was diluted with ethyl acetate and the organic phase washed sequentially with an aqueous solution of 1M sodium hydroxide and water, then dried. The organic extract was then filtered, and the filtrate concentrated to yield a crude solid which was purified by chromatography (DCM/MeOH 20:1), to yield 389

mg of a white powder (75 %). ^{1}H-NMR (400 MHz, $CDCl_3$): δ 7.81 (d, 2H, *J* = 9.0 Hz, H-4 and H-9 of naphthyl), 7.77 (d, 1H, *J* = 8.0 Hz, H-6 of naphthyl), 7.47 (m, 1H, H-7 of naphthyl), 7.44 (d, 1H, *J* = 2.5 Hz, H-1 of naphthyl), 7.38 (m, 2H, H-8, H-5 of naphthyl), 7.22 (dd, 1H, *J* = 2.5, 9.0 Hz, H-3 of naphthyl), 5.46 (d, 1H, *J* = 8.0 Hz, H-1), 5.39 (dd, 1H, *J* = 10.5, 10.5 Hz, H-3), 5.11 (dd, 1H, *J* = 10.5, 10.5 Hz, H-4), 4.36 (dd, 1H, *J* = 5.5, 12.5 Hz, H-6), 4.16 (M, 2H, H-2 and H-6), 4.11 (m, 1H, H-5), 2.07, 2.06, 2.04 (3×s, 9H, acetyl), 1.96 (s, 3H, acetamido). ^{13}C-NMR (100 MHz, $CDCl_3$): *δ* 20.8 (C-Me), 20.9 (C-Me), 20.9 (C-Me), 23.3 (C-Me), 59.0 (C-2), 61.9 (C-6), 71.0 (C-4), 74.6 (C-3), 78.0 (C-5), 106.7 (C-1), 121.0, 126.1, 126.5, 127.9, 128.0, 130.3, 137.4, 128.5, 148.1, 156.9 (C-naphthyl), 170.4 (C-Carbonyl), 171.4 (C-Carbonyl), 171.8 (C-Carbonyl), 172.5 (C-Carbonyl).



1-(2-Naphthyl) 2-acetamido-2-deoxy-β-D-glucopyranoside (43). To a solution of **44** (200 mg) in methanol-toluene (1:1) was added a catalytic amount of 0.5M sodium methoxide in methanol. The reaction mixture was stirred at room temperature for 0.5h and the progress of the reaction was monitored by TLC (DCM/MeOH 4:1). Upon completion of the reaction, the organic solution was concentrated, and the residue was purified by chromatography (DCM/MeOH 4:1), to yield 92 mg of a white powder (88%). ^{1}H-NMR (400 MHz, MeOD): *δ* 7.76 (m, 3H, H-4, H-9, H-6 of naphthyl), 7.42 (m, 2H, H-7, H-1 of naphthyl), 7.34 (m, 1H, H-8 of naphthyl), 7.18 (dd, 1H, *J* = 2.4, 9.2 Hz, H-3 of naphthyl), 5.19 (d, 1H, *J* = 8.4 Hz, H-1), 4.00 (dd, 1H, *J* = 8.4, 10.4Hz, H-2), 3.98 (dd, 1H, *J* = 2.4, 12Hz, H-6), 3.77 (1H, dd, *J* = 5.6, 12Hz, H-6), 3.63 (dd, 1H, *J* = 8.4, 10Hz, H-3), 3.58 (m, 1H, H-5), 3.48 (dd, 1H, *J* = 8.8,

9.6Hz, H-4), 1.99 (s, 3H, $NHCH_3$). ^{13}C-NMR (100 MHz, MeOD): δ 23.0 (C-Me), 57.4 (C-2), 62.6 (C-6), 71.9 (C-4), 75.9 (C-3), 78.4 (C-5), 101.0 (C-1), 111.9, 119.8, 125.3, 127.4, 128.2, 128.6, 130.4, 131.3, 135.9, 156.9 (C-naphthyl), 174.0 (C-Carbonyl). m/z (ESI) 371.1199 $[M+H+Na]^{2+}$, $C_{18}H_{22}NNaO_6$ requires 371.1345.

Enzymology. β-1,4-Galactosyltransferase (β-GalT) from bovine milk was either expressed in our own laboratory or obtained commercially from Sigma. For the expression of recombinant β-1,4-GalT we used the construct pET29b_b4GalT1Δ129 C342T, which was a generous gift from Dr Christelle Breton (Grenoble). The mutation C342T is known to improve stability and increase folding yield. Removal of the first 129 residues corresponds to the beginning of the catalytic domain. The protocol for renaturation of inclusion bodies and refolding was adapted from Qasba and co-workers.[37] For the biochemical phosphatase-coupled glycosyltransferse assays, we used a recently reported colorimetric protocol.[18] All assays were carried out in Nunc clear, flat-bottom 96-well polystyrene microplates. Assay wells typically contained $MnCl_2$, calf-intestinal phosphatase (CIP), chicken egg-white lysozyme (CEL), UDP-Gal donor and either GlcNAc or **43** as acceptor. To quantify the concentration of inorganic phosphate (P_i), malachite green reagents were added, and the absorbance was recorded at 620 nm on a BMG Labtech POLARstar Optima multiplate reader.

Data collection and analysis. A calibration curve (0–12.5 μM UDP, corresponding to 0-25 μM P_i) was constructed for each microplate by linear regression. The calibration curve was used to convert absorbance measurements at 620 nm in sample and control wells to [UDP] (μM). For each sample and control well, a corresponding background well (containing

identical components but no acceptor) was included, to account for non-specific hydrolysis of donor. Corrected absorbance values for each well were obtained by subtracting the corresponding background reading from the absorbance of the respective sample or control well. The calculated concentration of UDP was plotted against concentration of acceptor (for substrate assay or control assay) or incubation time (for time-dependent assays). Averages and standard deviations were calculated in Microsoft Excel.

LC/MS experiment. The standard assay mixtures contained UDP-Gal (500 μM), compound **43** (500 μM), β-1,4-GalT (50 or 200 μM), phosphatase (20 μM, in assay 4) in 13 mM HEPES buffer (pH = 7). Mixtures were incubated at 30 ºC for 1h. Reactions were stopped by the addition of the same volume of methanol. Then mixtures were centrifuged for 15min at 1000 rpm. The supernatants were used for LC/MS analysis directly. For LC/MS analysis: HPLC system with a detection set at 214 nm and coupled with column (Agilent Eclipse XDB-C8 4.6×150 mm) with a flow rate of 0.5 mL/min. The mobile phase comprised a mixture of water (0.1 % formic acid) and methanol. The utilized gradient was illustrated in **Table 4.2**. A Advion Compact Mass Spectrometer (CMS) was coupled with HPLC for mass detecting.

4.6. References

1. Wagner, G. K.; Pesnot, T., Glycosyltransferases and their assays. *Chembiochem* **2010,** *11*, 1939-49.

2. Palcic, M. M.; Sujino, K., Assays for Glycosyltransferases. *Trends Glycosci Glyc* **2001,** *13*, 361-370.

3. Palcic, M. M.; Pierce, M.; Hindsgaul, O., [14] Synthetic neoglycoconjugates in glycosyltransferase assay and purification. In *Methods in Enzymology*, Academic Press1994; Vol. Volume 247, pp 215-227.

4. Bailly, P.; Cartron, J.-P., Characterization and specific assay for a galactoside β-3-galactosyltransferase of human kidney. *Eur J Biochem***1988,** *173*, 417-422.

5. Yeoh, K. K.; Butters, T. D.; Wilkinson, B. L.; Fairbanks, A. J., Probing replacement of pyrophosphate via click chemistry; synthesis of UDP-sugar analogues as potential glycosyl transferase inhibitors. *Carbohydr Res* **2009,** *344*, 586-591.

6. Basu, M.; De, T.; Das, K. K.; Kyle, J. W.; Chon, H.-c.; Schaeper, R. J.; Basu, S., [51] Glycolipids. In *Methods in Enzymology*, Academic Press1987; Vol. Volume 138, pp 575-607.

7. Ahsen, O. v.; Voigtmann, U.; Klotz, M.; Nifantiev, N.; Schottelius, A.; Ernst, A.; Müller-Tiemann, B.; Parczyk, K., A miniaturized high-throughput screening assay for fucosyltransferase VII. *Anal Biochem* **2008,** *372*, 96-105.

8. Hayashi, Y.; Horibata, Y.; Sakaguchi, K.; Okino, N.; Ito, M., A sensitive and reproducible assay to measure the activity of glucosylceramide synthase and lactosylceramide synthase using HPLC and fluorescent substrates. *Anal Biochem* **2005,** *345*, 181-186.

9. Monegal, A.; Pinyol, R.; Planas, A., Capillary electrophoresis method for the enzymatic assay of galactosyltransferases with postreaction derivatization. *Anal Biochem* **2005,** *346*, 115-123.

10. Wu, J.; Takayama, S.; Wong, C.-H.; Siuzdak, G., Quantitative electrospray mass spectrometry for the rapid assay of enzyme inhibitors. *Chem Biol* **1997,** *4*, 653-657.

11. Yang, M.; Brazier, M.; Edwards, R.; Davis, B. G., High-Throughput Mass-Spectrometry Monitoring for Multisubstrate Enzymes: Determining the Kinetic Parameters and Catalytic Activities of Glycosyltransferases. *ChemBioChem* **2005,** *6*, 346-357.

12. Anumula, K. R., New high-performance liquid chromatography assay for glycosyltransferases based on derivatization with anthranilic acid and fluorescence detection. *Glycobiology* **2012,** *22*, 912-917.

13. Lee, K. B.; Desai, U. R.; Palcic, M. M.; Hindsgaul, O.; Linhardt, R. J., An electrophoresis-based assay for glycosyltransferase activity. *Anal Biochem* **1992,** *205*, 108-114.

14. Kanie, Y.; Kirsch, A.; Kanie, O.; Wong, C.-H., Enzymatic Assay of Galactosyltransferase by Capillary Electrophoresis. *Anal Biochem* **1998,** *263*, 240-245.

15. Schachter, H.; Brockhausen, I.; Hull, E., [30] High-performance liquid chromatography assays for N-acetylglucosaminyltransferases involved in N- and O-glycan synthesis. In *Methods in Enzymology*, Academic Press1989; Vol. Volume 179, pp 351-397.

16. Gosselin, S.; Alhussaini, M.; Streiff, M. B.; Takabayashi, K.; Palcic, M. M., A Continuous Spectrophotometric Assay for Glycosyltransferases. *Anal Biochem* **1994,** *220*, 92-97.

17. Wu, Z. L.; Ethen, C. M.; Prather, B.; Machacek, M.; Jiang, W., Universal phosphatase-coupled glycosyltransferase assay. *Glycobiology* **2011,** *21*, 727-33.

18. Tedaldi, L.; Evitt, A.; Goos, N.; Jiang, J.; Wagner, G. K., A practical glycosyltransferase assay for the identification of new inhibitor chemotypes. *Medchemcomm* **2014,** *5*, 1193-1201.

19. Wongkongkatep, J.; Miyahara, Y.; Ojida, A.; Hamachi, I., Label-Free, Real-Time Glycosyltransferase Assay Based on a Fluorescent Artificial Chemosensor. *Angew Chem Int Edit* **2006,** *45*, 665-668.

20. Deng, C.; Chen, R. R., A pH-sensitive assay for galactosyltransferase. *Anal Biochem* **2004,** *330*, 219-226.

21. Hang, H. C.; Yu, C.; Ten Hagen, K. G.; Tian, E.; Winans, K. A.; Tabak, L. A.; Bertozzi, C. R., Small Molecule Inhibitors of Mucin-Type O-Linked Glycosylation from a Uridine-Based Library. *Chem Biol* **2004,** *11*, 337-345.

22. Oubihi, M.; Kitajima, K.; Kobayashi, K.; Adachi, T.; Aoki, N.; Matsuda, T., Development of an Enzyme-Linked Immunosorbent Assay-Based Method for Measuring Galactosyltransferase Activity Using a Synthetic Glycopolymer Acceptor Substrate. *Anal Biochem* **1998,** *257*, 169-175.

23. Yan, L. Y.; Smith, D. F.; Cummings, R. D., Determination of GDP-Fuc:Galβ1-4GlcNAc-R (Fuc to GlcNAc) α1,3 Fucosyltransferase Activity by a Solid-Phase Method. *Anal Biochem* **1994,** *223*, 111-118.

24. Soltero-Higgin, M.; Carlson, E. E.; Phillips, J. H.; Kiessling, L. L., Identification of Inhibitors for UDP-Galactopyranose Mutase. *J Am Chem Soc***2004,** *126*, 10532-10533.

25. Helm, J. S.; Hu, Y.; Chen, L.; Gross, B.; Walker, S., Identification of Active-Site

Inhibitors of MurG Using a Generalizable, High-Throughput Glycosyltransferase Screen. *J Am Chem Soc* **2003,** *125*, 11168-11169.

26. Berndl, S.; Herzig, N.; Kele, P.; Lachmann, D.; Li, X.; Wolfbeis, O. S.; Wagenknecht, H.-A., Comparison of a Nucleosidic vs Non-Nucleosidic Postsynthetic "Click" Modification of DNA with Base-Labile Fluorescent Probes. *Bioconj Chem* **2009,** *20*, 558-564.

27. Helm, J. S.; Hu, Y.; Chen, L.; Gross, B.; Walker, S., Identification of active-site inhibitors of MurG using a generalizable, high-throughput glycosyltransferase screen. *J Am Chem Soc* **2003,** *125*, 11168-9.

28. Gross, B. J.; Swoboda, J. G.; Walker, S., A Strategy to Discover Inhibitors of O-Linked Glycosylation. *J Am Chem Soc* **2008,** *130*, 440-441.

29. Jares-Erijman, E. A.; Jovin, T. M., FRET imaging. *Nat Biotech* **2003,** *21*, 1387-1395.

30. Brockhausen, I.; Benn, M.; Bhat, S.; Marone, S.; Riley, J. G.; Montoya-Peleaz, P.; Vlahakis, J. Z.; Paulsen, H.; Schutzbach, J. S.; Szarek, W. A., UDP-Gal: GlcNAc-R β1,4-galactosyltransferase—a target enzyme for drug design. Acceptor specificity and inhibition of the enzyme. *Glycoconj J* **2006,** *23*, 525-541.

31. Gao, Y.; Lazar, C.; Szarek, W. A.; Brockhausen, I., Specificity of β1,4-galactosyltransferase inhibition by 2-naphthyl 2-butanamido-2-deoxy-1-thio-β-D-glucopyranoside. *Glycoconj J* **2010,** *27*, 673-684.

32. Sasaki, K.; Nishida, Y.; Kambara, M.; Uzawa, H.; Takahashi, T.; Suzuki, T.; Suzuki, Y.; Kobayashi, K., Design of N-acetyl-6-sulfo-beta-D-glucosaminide-based inhibitors of influenza virus sialidase. *Bioorgan Med Chem* **2004,** *12*, 1367-1375.

33. Sasaki, K.; Nishida, Y.; Kambara, M.; Uzawa, H.; Takahashi, T.; Suzuki, T.; Suzuki, Y.; Kobayashi, K., Design of N-acetyl-6-sulfo-β-d-glucosaminide-based inhibitors of influenza virus sialidase. *Bioorgan Med Chem* **2004,** *12*, 1367-1375.

34. Kajihara, Y.; Kodama, H.; Endo, T.; Hashimoto, H., Novel features of acceptor recognition by β-(1→4)-galactosyltransferase. *Carbohydr Res* **1998,** *306*, 361-378.

35. Rowland, A.; Knights, K. M.; Mackenzie, P. I.; Miners, J. O., The "albumin effect" and drug glucuronidation: bovine serum albumin and fatty acid-free human serum albumin enhance the glucuronidation of UDP-glucuronosyltransferase (UGT) 1A9 substrates but not UGT1A1 and UGT1A6 activities. *Drug metab dispo: the biological fate of chemicals* **2008,** *36*, 1056-62.

36. Chung, S. J.; Takayama, S.; Wong, C.-H., Acceptor substrate-based selective inhibition of galactosyltransferases. *Bioorg Med Chem Lett* **1998,** *8*, 3359-3364.

37. Ramakrishnan, B.; Shah, P. S.; Qasba, P. K., α-Lactalbumin (LA) Stimulates Milk β-1,4-Galactosyltransferase I (β4Gal-T1) to Transfer Glucose from UDP-glucose to N-Acetylglucosamine: CRYSTAL STRUCTURE OF β4Gal-T1·LA COMPLEX WITH UDP-Glc. *J Biol Chem* **2001,** *276*, 37665-37671.

5. Summary and outlook

During this project, different strategies for the development of uncharged inhibitors of β-1,4-GalT were explored. Based on the previously reported GalT inhibitor 5-FT UDP-Gal, the approach of pyrophosphate esterification of 5-FT UDP-Gal was attempted. This direct approach to mask the negative charge of the sugar nucleotide was unsuccessful due to synthetic challenges. It is hypothesised that the hydroxyl group in galactose ring could attack the pyrophosphate bond, facilitating the decomposition of the sugar nucleotide.

Next, nucleoside-based derivatives of 5-FT UDP-Gal were developed by formally removing the pyrophosphate and galactose moieties and optimising the substituent in the 5-position of uracil. Different aryl and heterocyclic substituents were introduced via microwave Suzuki cross-coupling reaction. One of the target compounds, 5-FT uridine, was utilized to further extend structural diversity through subsequent reductive amination with different amino acids. Amongst this series of nucleoside derivatives, compounds containing an indole moiety exhibited inhibitory activity towards β-1,4-GalTs. In order to investigate the respective contribution of the 5-substituent and the pyrophosphate and galactose group to enzyme inhibition, the corresponding sugar nucleotide derivatives were synthesized and assessed in biochemical assay. The target selectivity of active nucleosides and their corresponding sugar nucleotides was also investigated. One of the uncharged nucleoside inhibitors was used successfully in the cell assay to reduce the expression of the cell surface glycoprotein PSGL-1 in human monocytes. This application illustrates the potential of these novel, nucleoside-based β-1,4-GalT inhibitors as tool compounds for chemical biology and drug discovery, for example to study PSGL-1 medaited cell adhesion. They may also be potential inhibitor candidates for the development of anti-inflammatory agents.

In order to overcome the loss of activity of nucleoside derivatives due to the removal of pyrophosphate and galactose moieties, a DCC strategy was employed to identify suitable mimics of the pyrophosphate and sugar groups. The reversible reaction between an aldehyde fragment based on one of the most active nucleoside derivatives and a collection of hydrazines/hydrazides was used. In the presence of β-1,4-GalT, the amplification of an individual library member was observed by HPLC. The selected compound was synthesized and assessed in the biochemical assay. The synthesis of corresponding derivatives, the amide analogues and unsubstituted amine analogue, was carried out and their inhibitory activities were measured. Based on the SAR analysis of these compounds, it is suggested that the exact nature of the linkage between the nucleoside and hydrazide moieties has a critical effect on inhibition. Compared with the hydrazine substituents, nucleoside, especially the 5-position indole motif, appeared to be the dominant factor for inhibition. In future work, the effect of these new β-1,4-GalT inhibitors on the cell surface levels of PSGL-1 on human monocytes will be investigated.

A previously reported β-1,4-GalT inhibitor based on the GlcNAc acceptor substrate was selected as the positive control for the DCC experiments. Unexpectedly, this compound was found to behave as an enzyme substrate in the colorimetric assay used in this study. This result differed from previously reported findings in a radiochemical assay. A series of experiments was carried out to explain this discrepancy, including the detection of the glycosylated product by LC/MS. The phosphatase-coupled assay and LC/MS experiments confirmed that the presence of phosphatase could facilitate the formation of glycosylated products by the hydrolysis of nucleoside diphosphate.

After this study, it is unknown how the nucleoside derivatives bind to the catalytic site of GalTs, especially the interaction between the indole motif and enzyme. Thus, more

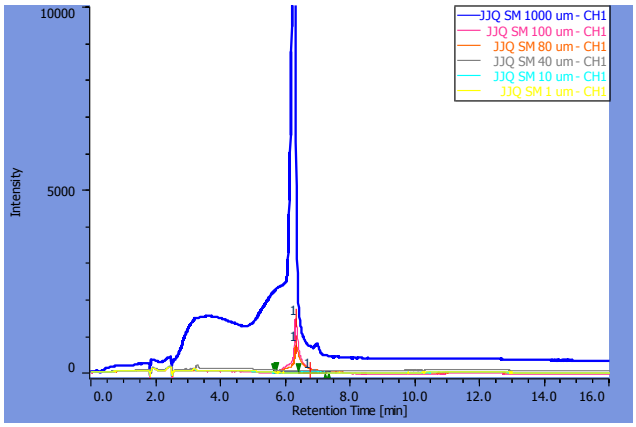
nucleoside derivatives with 5-substituent are required for SAR analysis in the future study. A crystal structure of the complex of enzyme and nucleoside analogue also would provide useful information about binding, which can help the further development of inhibitors towards GalTs. As the introduction of hydrazine as pyrophosphate and sugar moieties did not give additional activity, a detailed docking study is required for the investigation of binding between hydrazine moiety and enzyme active site. It will provide useful information for the modification of linkages for nucleoside and hydrazine, which is essential for the development of GalTs inhibitors.

6. Appendix

In this section, the chromatograms of HPLC results of chapter 3 was shown.

In order to measure the relationship between the compound concentration and UV detection of HPLC (**Figure A1**), different volumes of aldehyde **29** solution (300 μM) was applied to the HPLC using a JASCO MD-2018 with the detection at 254 nm. The **29** solution (1, 20, 40, 80, 100, 1000 μL) were separately injected on a reverse phase column (Luna 5u C8(2) 100A, new column 150 ×4.6 mm) and eluted at 0.5 mL/min. The mobile phase comprised a mixture of water (buffer A) and acetonitrile (buffer B) and the gradient A (**Table 12**) was utilized. The volume of 3.1 solution was plotted against the UV-absorbance intensity.

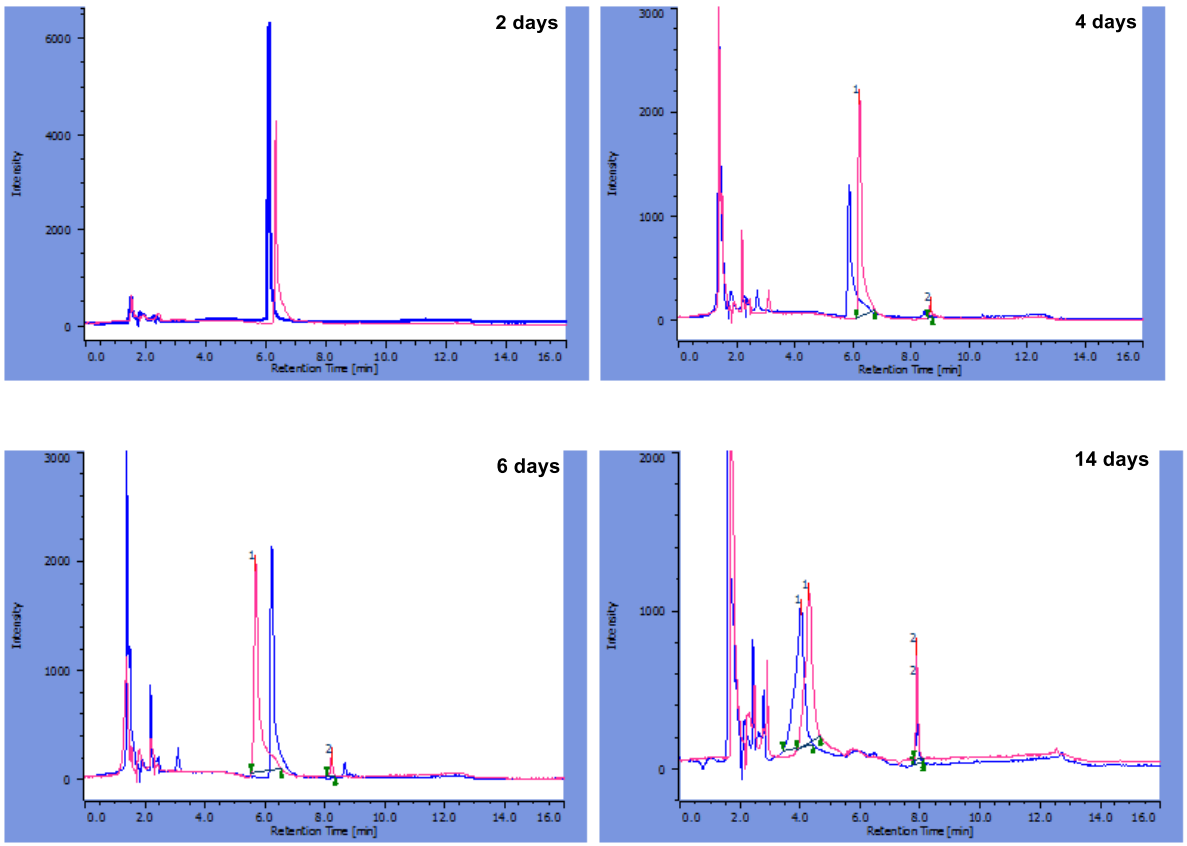
Figure A1 Chromatogram of the relation between amount of compound **29** and UV absorbance



DCC experiment was carried out with aldehyde **29** and **H-1**. Solutions of **H-1**, indole-uridine aldehyde **29** and cyanoborohydride were introduced into 0.5 mL eppendorf containing a solution of β1,4GalT in MOPS buffer. (Buffer: 50 mM MOPS, 20 MnCl2, pH 7.4) The final conc. of each reagents: **H-1** (300 μL), **29** (300 μL), $NaBH_3CN$ (5 mM), enzyme (0.69

mg/mL). Total volume of each eppendorf is 300 uL. Then incubated in 25 ºC and monitored by HPLC of a JASCO MD-2018 with the detection at 254 nm. 100 μL solution was injected on a reverse phase column (Luna 5u C8(2) 100A, new column 150 ×4.6 mm) and eluted at 0.5 mL/min. The mobile phase comprised a mixture of water (buffer A) and acetonitrile (buffer B) and the gradient A (**Table 12**) was utilized.

Figure A2 Chromatograms of reaction between **29** and **H-1** analysis in 2d, 4d, 6d, 14d. Reversible reaction incubated in the absence of enzyme (blue curve); Reversible reaction incubated in the presence of enzyme (pink curve).



DCC experiment was carried out with aldehyde **29** and **H-9**. Solutions of **H-9**, indole-uridine aldehyde **29** and cyanoborohydride were introduced into 0.5 mL eppendorf containing a solution of β1,4GalT in MOPS buffer. (Buffer: 50 mM MOPS, 20 MnCl2, pH 7.4) The final conc. of each reagents: **H-1** (300 μL), **29** (300 μL), $NaBH_3CN$ (5 mM), enzyme (0.69 mg/mL). Total volume of each eppendorf is 300 uL. Then incubated in 25 ºC and monitored

by HPLC of a JASCO MD-2018 with the detection at 254 nm. 100 μL solution was injected on a reverse phase column (Luna 5u C8(2) 100A, new column 150 ×4.6 mm) and eluted at 0.5 mL/min. The mobile phase comprised a mixture of water (buffer A) and acetonitrile (buffer B) and the gradient B (**Table 13**) was utilized.

Figure A3 Chromatograms of reaction between **29** and **H-9** analysis in 2d, 4d, 8d; Reversible reaction incubated in the absence of enzyme (blue curve); Reversible reaction incubated in the presence of enzyme (pink curve).